THE JOURNAL
OF
CHEMICAL PHYSICS

NUMBER 10 AUGUST, 1942 VOLUME 8

Gas Imperfections Determined from the Heat of Vaporization and Vapor Pressure

CHARLES F. CURTISS AND JOSEPH O. HIRSCHFELDER
Department of Chemistry, University of Wisconsin, Madison, Wisconsin
(Received May 8, 1942)

The imperfections of many gases may be obtained by combining available heats of vaporization and vapor pressure data with the Clausius-Clapeyron equation. The accuracy of this method in some cases is comparable to that of good direct P.V.T. measurements and therefore may be useful in preliminary investigations. If P is the vapor pressure, T is the absolute temperature, ΔH is the heat of vaporization, V and V_l are the molal volumes of the vapor and liquid phases:

$$\frac{PV}{RT}=\frac{-\Delta H}{R(1-V_l/V)[d\ln P/d(1/T)]}=1+\frac{\beta}{V}.$$

This equation determines the gas imperfection β. The value of β calculated in this manner is compared with experimental second virial coefficients for H_2O, NH_3, H_2, O_2, A, CH_4, propane, *n*-butane, *n*-heptane, ethene, propene, CH_3Cl, CCl_3F, CH_3OH, C_6H_6, and Hg. Values of β are determined for HI, HBr, HCl, CH_3NH_2, $(CH_3)_2NH$, HCN, isobutane, tetramethylmethane, 2,2,4, trimethylpentane, C_2N_2, COS, PH_3, H_2S, CH_3Br, and C_2H_5OH. Evidence of hydrogen bond formation is shown for CH_3OH, C_2H_5OH, CH_3NH_2, and $(CH_3)_2NH$. It is shown that HI and HCN contain a considerable fraction of instantaneously formed double molecules. The gas imperfection of benzene is perfectly normal. The number of instantaneously formed double molecules is estimated for all of the above molecular species.

THE imperfections of many gases may be obtained by using the Clausius-Clapeyron equation together with available experimental heats of vaporization and vapor pressure data. The accuracy in some cases is comparable to that of good direct P.V.T. measurements. In this way we can learn a great deal about the degree of association and the nature of the intermolecular forces in gases for which no direct P.V.T. measurements have ever been made. However, any conclusions based on this indirect method should be verified in other ways before they are regarded as scientific truth.

The Clausius-Clapeyron equation is usually written:

$$dP/dT=\Delta H/T(V-V_l). \qquad (1)$$

Here P is the vapor pressure, T is the absolute temperature at which the boiling takes place, ΔH is the heat of vaporization, and V and V_l are the molal volumes of the vapor and the liquid, respectively. This equation may be arranged into the form:

$$\frac{PV}{RT}=\frac{-\Delta H}{R(1-V_l/V)(d\ln P/d(1/T)]}. \qquad (2)$$

In elementary treatments the vapor is usually assumed to be ideal so that $PV/RT=1$ and the volume of the liquid is neglected in comparison with the volume of the vapor. When these assumptions are made, it is easy to integrate Eq. (2) to express the heat of vaporization directly in terms of the vapor pressure at two different temperatures. The values for the heats of vaporization obtained in this manner with good vapor pressure measurements are usually in error by around six percent. In this paper we

are attributing this deviation to the imperfection of the gas and making it serve a useful purpose. From the experimental measurements of vapor pressure we can calculate $d \ln P/d(1/T)$. If in addition we know ΔH and V_l/V, the right-hand side of the equation is completely determined and we are left with an experimental value of PV/RT.

It is convenient to express the imperfection of the gas in terms of a function, β, defined by the equation:

$$PV/RT=1+\beta/V. \tag{3}$$

We have tabulated the observed values of β for a number of substances. Large negative values correspond to association, and we can estimate the percentage of double molecules present.

The equation of state of any gas may be written in the virial form:

$$PV/RT=1+B(T)/V+C(T)/V^2+\cdots, \tag{4}$$

where $B(T)$, $C(T)$, $\cdots$, are called the second, third, $\cdots$, virial coefficients and are functions of the temperature but not of the pressure. In ordinary P.V.T. measurements one can determine the virials separately by varying the pressure while keeping the temperature constant. This is not possible in our method since we can only consider one pressure for any particular temperature. However, most of the experimental vapor pressure measurements are made at pressures sufficiently low that the contribution of third and higher virials is negligible for a normal gas, and for these we can show that $\beta=B(T)$. For associated gases where β is exceptionally large it is probable that the third and higher virials appreciably affect the value of β.

The argument that $\beta=B(T)$ for normal gases measured at a vapor pressure of around one atmosphere or less depends on the theorem of corresponding states. If we divide the second virial coefficient by the critical volume V_c we get a reduced second virial coefficient $B_r(T/T_c)$, which should be a universal function of the reduced temperature T/T_c. Similarly, we could define reduced third and higher virial coefficients. Hirschfelder, McClure, and Weeks[1] have shown that for a number of normal gases the experimental second virial coefficient agrees reasonably well with that which would apply if energy of interaction of the molecules satisfied a Lennard-Jones potential:

$$E(r)=4E_m[(r_0/r)^{12}-(r_0/r)^6], \tag{5}$$

where E_m, the maximum energy of interaction between pairs of molecules on collision, is

$$E_m=0.73T_ck \tag{6}$$

and k is the Boltzmann constant. Now we can make use of the work of Montroll and Mayer[2] who express the third and fourth virial coefficients for the Lennard-Jones gases in the form:

$$\begin{aligned} C(T)&=-8/3(B(T))^2\nu_2 \\ D(T)&=6(B(T))^3\nu_3. \end{aligned} \tag{7}$$

Here ν_2 and ν_3 are known functions of E_m/kT. The normal boiling point usually corresponds to values of T approximately $0.6T_c$ or

$$E_m/kT=0.73T_ck/0.6T_ck=1.2. \tag{8}$$

For this argument Montroll and Mayer computed that $\nu_2=0.019$ and $\nu_3=1.13$. For normal gases, in this temperature range, $|B/V|$ is always less than 0.05 so that

$$\begin{aligned} |C/V^2|&<|(8/3)(0.05)^2(0.019)|=0.0002 \\ \text{and}\quad |D/V^3|&<|(6)(0.05)^3(1.15)|=0.001. \end{aligned} \tag{9}$$

Therefore, in identifying β with $B(T)$ we are neglecting terms in PV/RT of the order of 0.001 or making an error in $B(T)$ of the order of $22{,}000\times0.001$ of 22 cc per mole. This is about the limit of the accuracy of β which can be attained from experimental considerations.

The ratio of double to single molecules present in the gas at any instant, N_2/N_1 is estimated by the method developed by Hirschfelder, McClure, and Weeks.[1] In place of $B(T)$ we take β and for the van der Waals b for the definitional domain of the double molecules we take $1.75V_c$. Then the equilibrium constant for the formation of double molecules is:

$$K=\frac{N_2V}{N_1^2}=1.75V_c-\beta.$$

Since we are usually dealing with a small degree

[1] J. O. Hirschfelder, F. T. McClure, I. S. Weeks, J. Chem. Phys. **10**, 201 (1942).

[2] According to a private communication from J. Mayer. The development of these higher virial coefficients was subsequently published by E. W. Montroll and J. E. Mayer, J. Chem. Phys. **9**, 626 (1941).

TABLE I.

	T°K	$-\beta$ calc.	$-B(T)$ obs.	
1. H_2O	373.16	460	452	
	423.16	287	284	
	473.16	196	197	
	523.16	139	146	
	573.16	97	112	
	623.16	58	89	
2. NH_3	233.16	516	662[a]	561[b]
	239.68	438	593	507
	243.16	443	560	467
	273.16	329	367	319
	293.16	265	290	250
	313.16	217	237	207
3. H_2	18.2	164[c]	210	
	20.4	72[d]	170	
O_2	90.13	206[e]	230	
A	87.29	244[f]	220	
CH_4	99.54	4090[g]	340	
	111.55	203[g]	290	
4. propane	231.04	658[h]	700	
n-butane	272.66	904[i]	840	
n-heptane	371.51	1300[j]	1500	
ethene	169.40	428[k]	430	
propene	225.35	575[l]	650	
CH_3Cl	248.94	527[m]	680	
$CFCl_3$	290.40	831[n]	620	
5. CH_3OH	298.16	4190[o]	2280	
	323.16	2130	1370	
	337.86	1480	1030	
	363.16	905	640	
	383.16	655	440	
Benzene	323.16	1390[p]	1610	
	353.16	885	1160	
	383.16	746	850	
6. Hg	630.0	2050[q]	250	

[a] Reference 5.
[b] Reference 9.
[c] Martinez and Onnes, Comm. Phys. Lab. Univ. Leiden **156b** (1922); Bichowsky and Rossini, *Thermochemistry of Chemical Substances* (Reinhold, 1936).
[d] Martinez and Onnes, Comm. Phys. Lab. Univ. Leiden **156b** (1922); Eucken, Ber. d. Deut. Phys. Ges. **18**, 4 (1916).
[e] Cath, Comm. Phys. Lab. Univ. Leiden **152d** (1918); Giauque and Johnston, J. Am. Chem. Soc. **51**, 2300 (1929).
[f] Born, Ann. d. Physik **69**, 473 (1922); Frank and Clusius, Zeits. f. physik. Chemie **42**, 395 (1939).
[g] Henning and Stock, Zeits. f. Physik **4**, 226 (1921); Frank and Clusius, Zeits. f. physik. Chemie **42**, 395 (1939).
[h] Kemp and Egan, J. Am. Chem. Soc. **60**, 1523 (1938).
[i] Aston and Messerly, J. Am. Chem. Soc. **62**, 1917 (1940).
[j] Smith, J. Research Nat. Bur. Stand. **24**, 229 (1940); Pitzer, J. Am. Chem. Soc. **62**, 1224 (1940).
[k] Egan and Kemp, J. Am. Chem. Soc. **59**, 1265 (1937).
[l] Powell and Giauque, J. Am. Chem. Soc. **61**, 2367 (1939).
[m] Messerly and Aston, J. Am. Chem. Soc. **62**, 886 (1940).
[n] Osborne, Garner, Doescher, and Yost, J. Am. Chem. Soc. **63**, 3496 (1941).
[o] Young, Proc. Roy. Soc. Dublin **12**, 374 (1910); Fiock, Ginnings, and Holton, J. Research Nat. Bur. Stand. **6**, 881 (1931).
[p] Dejardin, Ann. de Physique **12**, 253 (1919); Fiock, Ginnings, and Holton, J. Research Nat. Bur. Stand. **6**, 881 (1931).
[q] Menzies, Zeits. f. physik. Chemie **130**, 90 (1927); W. Kurbatoff, *ibid.* **43**, 104 (1903).

of dissociation, to the first approximation:

$$V=\frac{N_1RT}{P}.$$

So that:

$$\frac{N_2}{N_1}=\frac{PV_c}{RT}\left[1.75-\frac{\beta}{V_c}\right].$$

I. EXPERIMENTAL VERIFICATION OF THE METHOD

Before relying on such an indirect method of determining the imperfections of gases it is obligatory to make a rather extensive comparison between our values of β and those of $B(T)$ calculated from direct P.V.T. measurements. In the cases of NH_3 and H_2O sufficient data are available to permit us to make the comparison over a considerable temperature range. For H_2, O_2, A, CH_4, CH_3Cl, CCl_3F, propane, *n*-butane, *h*-heptane, ethene, and propene we checked our values at a temperature in the neighborhood of the normal boiling point. In Table I, β is compared with the second virial coefficients obtained directly. Most of the values agree within the accuracy of the experimental data considered; only a few values are poor.

1. H_2O

Our values of β for steam were determined from the empirical equation of Keenan and Keyes[3] for the vapor pressure as a function of the temperature and from the heats of vaporization of water determined by Osborne, Stimson, and Ginnings.[4] It was necessary to differentiate the vapor pressure equation to obtain $d \ln p/d(1/T)$. This is probably the largest single source of error in our method. The values of β were compared to those of $B(T)$ calculated from the equation of F. G. Keyes.[5] The experimental values for steam are exceptionally accurate so that it is not surprising that the agreement between β and $B(T)$ is good.

2. NH_3

Osborne and Van Dusen[6] determined the heats of vaporization of NH_3 at various temperatures.

[3] J. H. Keenan and F. G. Keyes, *Thermodynamical Properties of Steam* (John Wiley, 1936).
[4] Osborne, Stimson, and Ginnings, J. Research Nat. Bur. Stand. **23**, 261 (1939).
[5] F. G. Keyes, J. Am. Chem. Soc. **60**, 1761 (1938).
[6] N. S. Osborne and M. S. Van Dusen, J. Am. Chem. Soc. **40**, 1 (1918).

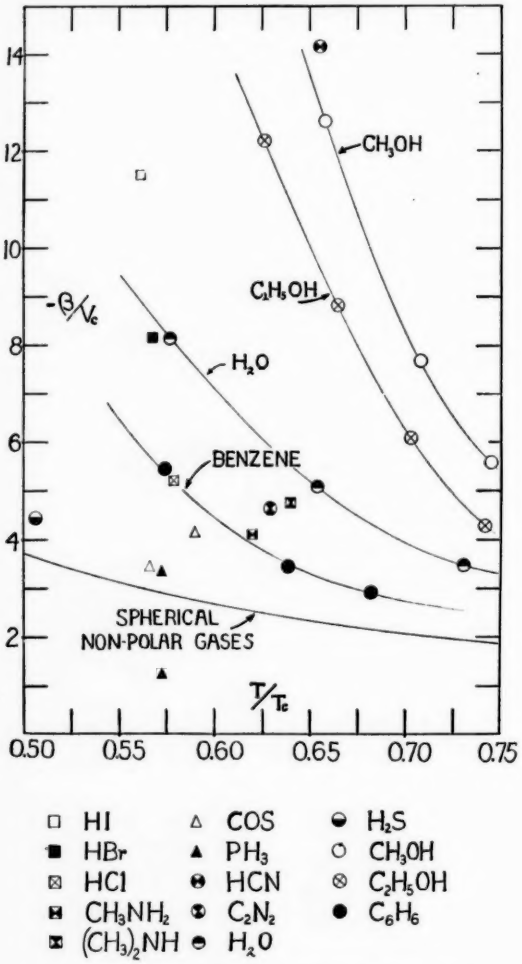


FIG. 1. Reduced second virial coefficients as a function of the reduced temperature.

Overstreet and Giauque[7] developed an equation for the vapor pressure of NH_3 in terms of the temperature, from 199°K to 242°K. They also determined the heat of vaporization at 239.68°K. This work was done in process of determining the entropy calorimetrically. Cragoe, Meyers, and Taylor[8] developed an equation for the vapor pressure from −78°C to 70°C. We calculated β from these data.

Keyes[5] also gives an equation for the virial coefficient of NH_3. This equation gives us values of $B(T)$ to compare with our values of β. In order to compare our accuracy with that of various direct P.V.T. determinations, we calculated the second virial coefficient from the previously accepted data for NH_3 as tabulated in a publication of the Bureau of Standards.[9] The differences between the two relatively accurate P.V.T. determinations of $B(T)$ is of the same order of magnitude as that between β and $B(T)$.

3. Simple Non-Polar Gases

The second virial coefficients for the simple non-polar gases were estimated on the assumption that the molecules obey a Lennard-Jones type of potential. For these calculations we used previously determined values of the molecular constants.[10] These values agree reasonably well with our values of β.

4. Hydrocarbons and Substituted Methanes

These were compared to values extrapolated from experimental data at high temperatures by means of equations developed recently.[1] The value of β for CH_4 at 99.54°K is obviously erroneous and probably due to small experimental errors.

5. Methyl Alcohol and Benzene

Eucken and Meyer[11] have measured directly the second virial coefficients for a number of organic substances. They included methyl alcohol and benzene. Their values are compared to ours. The values of $B(T)$ are calculated from their empirical equations. This amounts to a short extrapolation in both cases. At the higher temperatures the agreement is satisfactory, i.e., within 25 percent.

6. Hg

Our value is of doubtful accuracy since the heat of vaporization was determined by W. Kurbatoff in 1903 by a rather crude method. However it is interesting to note that this value is in line with the fact that mercury vapor contains a considerable fraction of double molecules. Actually mercury contains such a large fraction

[7] R. Overstreet and W. F. Giauque, J. Am. Chem. Soc. **59**, 254 (1937).
[8] Cragoe, Meyers, and Taylor, Bur. Stand. Sci. Papers **16**, 1 (1920).
[9] Circular of the Bureau of Standards, No. 142 (1925).
[10] J. O. Hirschfelder and W. E. Roseveare, J. Phys. Chem. **43**, 15 (1939).
[11] A. Eucken and L. Meyer, Zeits. f. physik. Chemie **5b**, 452 (1929).

of Hg_2 that its spectrum is well known and spectroscopic constants have been calculated for it. Hildebrand, Wakeham, and Boyd[12] determined the potential energy curve for Hg_2 from this spectroscopic data and Robert Bonner[13] has estimated from it the second virial coefficient as a function of temperature. The large deviation between β and the calculated $B(T)$ may or may not be significant. Further experimental work on the heat of vaporization and vapor pressure of mercury would be very desirable.

II. APPLICATION OF THE METHOD

Since this method appears to be valid whenever the experimental values of the vapor pressure and heats of vaporization are reasonably accurate, it is interesting to examine data for a number of different molecular species for which there are no direct P.V.T. data. In Fig. 1 we have plotted the values of $-\beta/V_c$ *versus* T/T_c for a number of substances. V_c and T_c are the critical volume and temperature, respectively. If the equation of corresponding states were valid, all of these points should lie on one curve. Actually the points show no such regularity. The reason for this is that the particular molecules studied here were chosen because of certain unusual properties, i.e., tendency to form hydrogen bonds, asymmetrical shape, large dipole moments, and the like. The reason for the deviations from the corresponding states equation is explained for the individual molecules. In Table II, our results are summarized and references given to all of the experimental data. We have included a column for μ^2/T_cV_c (where μ is the dipole moment) since the value of this reduced dipole moment should govern the deviation from the corresponding states equation due to dipole effects. The fact that this is not true for molecules forming hydrogen bonds shows that the formation of hydrogen bonds must be considered quite separately from normal dipole behavior. Undoubtedly many of our values of β are unreliable because of errors in the experimental heats of vaporization and vapor pressure, but it is amusing to see what implications could be drawn if we assume that all of our values are correct.

[12] Hildebrand, Wakeham, and Boyd, J. Chem. Phys. **7**, 1094 (1939).
[13] Robert Bonner, unpublished.

Where possible the critical constants were taken from the *International Critical Tables*. In a few cases where no data on the critical volume were available, we estimated the value from the critical temperature and pressure and a value of pv/RT at the critical conditions equal to that of a similar substance. The dipole moments are those recommended by Smyth.[14]

TABLE II.

	T°K	$-\beta$	T/T_c	$-B/V_c$	μ^2/T_cV_c ($\times 10^{-41}$)	N_2/N_1 (at 1 atmos.)
HI	237.75	1300[a]	.561	11.50	.30	.077
HBr	206.38	767[b]	.568	8.14	1.78	.055
HCl	188.07	446[c]	.579	5.19	3.80	.039
CH_3NH_2	266.84	518[d]	.620	4.11	2.79	.034
$(CH_3)_2NH$	280.04	868[e]	.640	4.74	1.15	.052
HCN	298.80	1910[f]	.654	14.16	11.4	.088
Isobutane	261.44	695[g]				
Tetramethyl-methane	278.59	1090[h]				
2,2,4, Trimethyl-pentane	372.33	1480[i]				
C_2N_2	251.95	695[j]	.628	4.65	0	.046
COS	214.03	470[k]	.566	3.45	.82	.040
	222.87	565[l]	.590	4.15		.044
PH_3	185.35	379[m]	.572	3.35	.83	.038
	185.72	143[n]	.573	1.26		.032
H_2S	188.7	410[o]	.505	4.45	2.63	.037
	212.77	1740[p]	.570	18.9		.109
CH_3Br	276.66	660[q]				
CH_3OH	298.16	4190[r]	.581	35.6	4.67	.180
	323.16	2130	.630	18.1		.088
	337.86	1480	.658	12.6		.061
	363.16	905	.708	7.68		.037
	383.16	655	.747	5.56		.027
C_2H_5OH	323.16	2040[r]	.626	12.2	3.35	.087
	343.16	1470	.665	8.82		.063
	363.16	1020	.703	6.09		.044
	383.16	715	.742	4.28		.032
Benzene	323.16	1390[s]	.575	5.43	0	.069
	353.16	885	.639	3.45		.046
	383.16	746	.682	2.90		.038

The first reference refers to the source of data on the vapor pressure of the material. The second refers to the source of data on the heat of vaporization. Where only one reference is given, this contains data on both.

[a] Henglein and Roth, Zeits. f. Physik **18**, 64 (1923); Giauque and Wiebe, J. Am. Chem. Soc. **51**, 1441 (1929).
[b] Henglein and Roth, Zeits. f. Physik **18**, 64 (1923); Giauque and Wiebe, J. Am. Chem. Soc. **50**, 2193 (1928).
[c] Henning and Stock, Zeits. f. Physik **4**, 226 (1921); Giauque and Wiebe, J. Am. Chem. Soc. **50**, 101 (1928).
[d] Aston, Siller, and Messerly, J. Am. Chem. Soc. **59**, 1743 (1937).
[e] Aston, Eidinoff, Forster, J. Am. Chem. Soc. **61**, 1539 (1939).
[f] Giauque and Ruehrwein, J. Am. Chem. Soc. **61**, 2626 (1939).
[g] Aston, Kennedy, and Schumann, J. Am. Chem. Soc. **62**, 2059 (1940).
[h] Aston and Messerly, J. Am. Chem. Soc. **58**, 2355 (1936).
[i] Smith, J. Research Nat. Bur. of Stand. **24** 229 (1940); Pitzer, J. Am. Chem. Soc. **62**, 1224 (1940).
[j] Ruehrwein and Giauque, J. Am. Chem. Soc. **61**, 2940 (1939).
[k] Kemp and Giauque, J. Am. Chem. Soc. **59**, 79 (1937); Frank and Clusius, Zeits. f. physik. Chemie **42**, 395 (1939).
[l] Kemp and Giauque, J. Am. Chem. Soc. **59**, 79 (1937).
[m] Stephenson and Giauque, J. Chem. Phys. **5**, 149 (1937).
[n] Henning and Stock, Zeits. f. Physik **4**, 226 (1921); Frank and Clusius, Zeits. f. physik. Chemie **42**, 395 (1939).
[o] Giauque and Blue, J. Am. Chem. Soc. **58**, 831 (1936); Frank and Clusius, Zeits. f. physik. Chemie **42**, 395 (1939).
[p] Giauque and Blue, J. Am. Chem. Soc. **58**, 831 (1936).
[q] Egan and Kemp, J. Am. Chem. Soc. **60**, 2097 (1938).
[r] Young, Proc. Roy. Soc. Dublin **12**, 374 (1910); Fiock, Ginnings, and Holton, J. Research Nat. Bur. Stand. **6**, 881 (1931).
[s] Dejardin, Ann. d. Physique **12**, 253 (1919); Fiock, Ginnings, and Holton, J. Research Nat. Bur. Stand. **6**, 881 (1931).

[14] C. P. Smyth, *Dielectric Constant* (Chemical Catalog Company, New York, 1931).

1. HI, HBr, HCl

For this series HI has by far the largest value of $-\beta/V_c$ in spite of the fact that its dipole moment is the smallest. Apparently HI is really made up of a considerable fraction of H_2I_2.

2. CH_3NH_2 and $(CH_3)_2NH$

Dimethylamine has a large value of $-\beta/V_c$ and seems to be largely associated. Methylamine has a somewhat smaller value of $-\beta$ but still contains a considerable fraction of dimers. These results are not surprising since the amines and particularly the diamines have a tendency to form hydrogen bonds. A straight dipole effect should give methyl amine the larger value of $-\beta/V_c$ so that its effect alone should give the opposite order. While the above results are suggestive, they are not conclusive since they might arise from experimental errors.

3. Methyl and Ethyl Alcohol

Both methyl and ethyl alcohols have very large values for $-\beta/V_c$ which would indicate that they contained a considerable fraction of dimers in the gas phase. These results are in keeping with the tendency for the alcohols to form hydrogen bonds. Considerable work has been done on the study of hydrogen bonds formed in liquid alcohol.

4. Benzene

The values of $-\beta/V_c$ for benzene seem perfectly normal and would agree with the corresponding states equation for the second virial coefficients of the hydrocarbons. Benzene molecules are shaped like plates and interfere with each other's rotation in the liquid. (This argument is used to explain why the benzene made from deuterium has a greater vapor pressure than ordinary benzene.) On this account, we should expect an unusually small number of double molecules. Direct P.V.T. determination of the second virial coefficient of benzene would be particularly useful for theoretical interpretations.

5. HCN

It appears that HCN gas contains an appreciable fraction of instantaneously formed double molecules. This agrees with the conclusion reached by Giauque and Ruehrwein[15] that HCN vapor contains straight chain polymers of indefinite length.

CONCLUSION

1. The Clausius-Clapeyron equation combined with accurate experimental values of the heats of vaporization and vapor pressures suffices to make reasonable estimates of the second virial coefficients for a number of substances for which no direct P.V.T. data is available.

2. The most interesting examples are those molecules which have a tendency to form hydrogen bonds. Our calculations for the fraction of dimers present in HI, dimethylamine, methyl and ethyl alcohol, and HCN should be substantiated or refuted by infra-red spectroscopic observations of the vapors.

3. This method is particularly suitable for use with those compounds whose absolute entropies are determined with great precision by calorimetrical methods and accurate heats of vaporization and vapor pressures are determined as by-products.

[15] W. F. Giauque and R. A. Ruehrwein, J. Am. Chem. Soc. **61**, 2626 (1939).

The Photolysis of Dimethylhydrazine[1]

WILLIAM L. KAY AND H. AUSTIN TAYLOR
Department of Chemistry, New York University, New York, New York

(Received June 11, 1942)

The ultraviolet absorption spectrum of dimethylhydrazine has been shown to exhibit a continuum extending from 2800A to beyond 2000A. Despite the absence of structure the continuum is probably experimental and indicates predissociation. The photolysis using various light sources, at several pressures and wave-lengths, has been studied. The over-all quantum yield is 0.3. The principal initial rupture of the molecule involves the scission of a hydrogen atom. Analytically three molecules of the hydrazine give one molecule each of H_2, N_2, CH_4, NH_3, CH_3NH_2, and $(CH_3N : CH_2)_2$. The low quantum yield is accounted for by addition of hydrogen to the radical produced in the initial split.

THE reaction of atomic hydrogen with azomethane was shown by Henkin and Taylor[2] to yield *s*-dimethylhydrazine alone at room temperature. At 110°C there was, in addition, a rupture producing methane and methylamine. Whether this rupture was solely of the azomethane, the dimethylhydrazine, or of both could not be clearly deduced. The photochemical decomposition of hydrazine has been studied but no work has been reported on *s*-dimethylhydrazine. The absorption spectrum and over-all quantum yield has therefore been measured at various wave-lengths.

EXPERIMENTAL

Materials

Dimethylhydrazine prepared from the dihydrochloride by treatment with pellets of potassium hydroxide was drawn through a drying column filled with fused barium oxide and condensed in a trap cooled with dry ice. The liquid was freed from dissolved gases by repeated evaporation and condensation under vacuum. The boiling point, determined by a micro method, was 80.0°C in agreement with that reported by West,[3] 80.5°C. The refractive index at 20°C using a sodium D lamp was 1.4204; West reports 1.4201. The density was 0.826 as against West's value 0.827.

Propylene used in the tests was taken from a commercial tank and repeatedly fractionated in vacuum from dry ice to liquid nitrogen temperatures. Analysis showed the presence of 5 percent propane and 94 percent unsaturate.

Uranium oxalate was prepared by cooling a hot solution of reagent quality uranium nitrate and oxalic acid. The resulting crystals were thoroughly washed, dried overnight in vacuum and then in air for three hours at 110°C. Oxalic acid and sodium oxalate were recrystallized several times, dried overnight in vacuum and finally in air at 105°C for three hours before use.

Apparatus

The reaction cell of fused quartz was a cylinder, 20-cm inside length and 2.5-cm inside diameter. It was joined through a ground glass joint to a Pyrex manifold evacuated by a mercury diffusion pump and oil pump. The cell was removed for cleaning prior to each experiment. The cell was connected to a manometer and a series of traps through which the products of the reaction could be withdrawn by a mercury piston and delivered to a gas buret for measurement of volume or for collection over mercury for subsequent analysis. A small removable trap between the cell and mercury piston was filled with yellow mercuric oxide prior to each experiment.

Radiation

Various light sources were used in the work:

Arc I: A "cold" mercury vapor lamp operated from a transformer drawing 110 volts in the

[1] Abstract from a thesis presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy at New York University, March, 1942.
[2] H. Henkin and H. A. Taylor, J. Chem. Phys. **8**, 1 (1940).
[3] W. West and R. B. Killingsworth, J. Chem. Phys. **6**, 1 (1938).

primary and delivering 15 volts at the secondary.
Arc II: A "hot" mercury arc by the Hanovia Company of the Alpine sunlamp type.
Arc III: A General Electric Type H4 lamp with the outer Pyrex envelope removed, operating from a transformer drawing 2 amp. in the primary at 120 volts.
Arc IV: A low pressure vapor lamp used for the quantum yield determinations. It was wound as a helical spiral of 10 turns with an internal diameter of 30 mm and operated from a 0.265-kva transformer.
Spark: Several runs were made using a condensed spark between electrodes of Zn, Cd, or Al. The power was furnished by a 10-kva transformer drawing 30 amp. at 250 volts and delivering 10,500 volts on the secondary. A Pyranoil capacitor of 0.0375 μf was placed parallel with the transformer. The electrodes were metal collars 10-cm diameter and 5 mm thick with V-shaped edges mounted on 6-cm disks carried by water-cooled shafts rotating at about 80 r.p.m. The electrodes ran at right angles to each other and the adjustable gap between them was about 4 mm. The set-up is similar to that used by Wiig and Kistiakowsky.[4]

Absorption Measurements

The absorption spectrum of dimethylhydrazine was determined using a Hilger E-2 spectrograph. The light source was a hydrogen discharge tube supplied by the Hanovia Company operating at about 2500 volts and 0.3 amp. Eastman 50 and ultraviolet sensitized plates were used. The intensities of the photographed spectra were measured on a Kipp recording microphotometer. Three absorption cells were used, of 10, 20, and 120-cm length.

Procedure

The system being thoroughly evacuated (10^{-6} mm or less on a McLeod gauge), the dimethylhydrazine, freshly pumped out at dry ice temperature, was allowed to warm up and fill the reaction cell to the desired pressure which was always less than the saturated vapor pressure at room temperature. Exposure to the arc or spark which had been permitted to reach a steady state while shielded was then made. Pressure changes during the run were followed on the manometer; the final pressure was read after the light had been turned off for several minutes. The products of reaction were drawn by the mercury piston through a trap in liquid nitrogen and their volume measured in the buret at atmospheric pressure. A second fractionation was made using melting sec-butyl chloride at −131°C and a third from dry ice at −78°C. The gases were collected over mercury and analyzed in a Blacet and Leighton micro gas analyzer.[5]

The gas drawn from dry ice was condensed in the yellow mercuric oxide trap to remove ammonia after its volume had been measured. The trap and its contents were allowed to warm to room temperature and again cooled several times until no further decrease in volume occurred. The residual gas was withdrawn and tested with an acid bead for complete absorption and taken to be methylamine.

The gas condensed in the dry ice trap was warmed to room temperature and returned to the reaction cell where a measurement of the pressure permitted the calculation of its volume. This was taken as the undecomposed dimethylhydrazine.

Quantum yield determinations were made using a uranyl oxalate actinometer, 0.01M uranyl oxalate and 0.05M oxalic acid.[6] Measurements were made using Arc IV, a fused quartz lens serving to focus the light through an iris diaphragm on the cell, and, since the lens only transmitted down to 2400A, it filtered out the 1850A radiation. The actinometer vessel was of fused quartz 4 cm in length with planar end plates 2.5 cm in diameter. The solution was vigorously stirred during photolysis by a platinum ribbon. A blank of the same solution was protected from light and titrated with permanganate at the same time as the photolyzed solution. Two runs were made with the actinometer placed behind the empty reaction cell and then behind the cell filled with dimethylhydrazine. The normal procedure, however, was to make a run

[4] E. O. Wiig and G. B. Kistiakowsky, J. Am. Chem. Soc. **54**, 1806 (1932).

[5] F. E. Blacet and P. A. Leighton, Ind. Eng. Chem., Anal. Ed. **3**, 266 (1931); **5**, 272 (1933); **6**, 344 (1934); **9**, 44 (1937).

[6] G. S. Forbes and L. J. Heidt, J. Am. Chem. Soc. **56**, 2363 (1934).

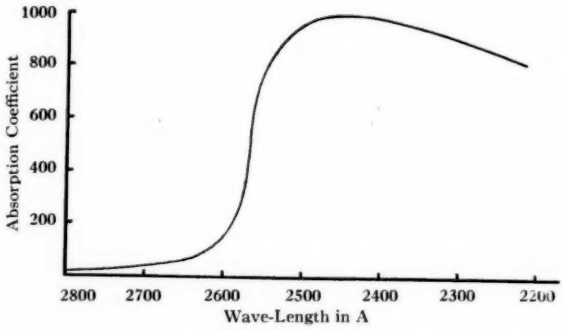


FIG. 1. Absorption coefficients of dimethylhydrazine.

with dimethylhydrazine and then replace the reaction cell by the actinometer vessel.

In focal isolation the full light from the zinc spark was focused by two crystal quartz lenses and two small diaphragms each at optimum position. The lenses were plano-convex and carried stops of black paper mounted centrally on their plane faces. A small quartz spectrograph was used to determine the correct focus for the desired lines. Spectrograms showed the zinc lines 2030, 2060, 2100, and 2140A could be isolated as a group.

In determining the absorption spectrum, exposures were first made for various times through the empty cell and then through the filled cell for the same times. This was repeated for various pressures of dimethylhydrazine. A copper spectrum was used for comparison.

RESULTS

Absorption Spectrum

The absorption spectrum of dimethylhydrazine was determined at pressures ranging from 0.1 to 50 mm in cells 10, 20, and 120 cm in length. The spectrum showed continuous absorption commencing at 2800A and continuing through the limit of the plate. Spectrograms taken using the higher resolution of the Hilger E-1 instrument showed no structure either on the plate or in the photometer tracings.

The absorption coefficient k, given by $I/I_0 = 10^{-kcl}$ where c is the concentration in moles per liter and l is the length in cm, was calculated using the characteristic curve for the plate determined from the photometer tracings of the exposures through the empty cell. In Fig. 1 data are presented for 4.2-mm pressure in the 120-cm cell. It is clear that at 2537A absorption will be effectively complete under the experimental conditions of the photolysis.

Decomposition Products

The products of photolysis were shown to consist of hydrogen, methane, nitrogen, ammonia, methylamine, and a liquid. The first three, not condensed by liquid nitrogen, were analyzed by combustion with oxygen. Hydrogen was determined in a separate sample using a cuprate bead. Nitrogen was obtained from the difference between the total sample and the hydrogen and methane found.

The volume of the sec-butyl chloride fraction was never more than 1/15 the volume of the liquid nitrogen fraction. Analyzed by combustion it consisted principally of methane although in runs using the hot arc placed close to the reaction cell, when the amount of dimethylhydrazine decomposed was large, small amounts of ethane were found. Introduction of a fuming sulfuric acid bead into this fraction before combustion caused a contraction of 3–4 percent. This might be caused by small amounts of ammonia being drawn over.

The presence of ammonia and methylamine in the fraction drawn from dry ice was established in several ways. In the first place Tanret[7] says that the ammonia reaction of Nessler's reagent is not given by methyl or ethylamine. Charitschkoff[8] claims that the amine hydrochlorides which are decomposed by aqueous potassium iodide give the reaction. It was found in synthetic methylamine-ammonia mixtures that with appreciable amounts of methylamine a precipitate resulted. If the solution is greatly diluted, however, it is possible to determine the ammonia colorimetrically.

An aqueous solution of the gases drawn from dry ice gave a yellow precipitate with Nessler's reagent. A solution of pure methylamine gave a much lighter colored precipitate which darkened on addition of ammonia and could be matched with the former. Dimethylamine gives a white precipitate which even on addition of large amounts of ammonia could not be made to

[7] C. J. Tanret, J. pharm. chim. **5**, 28, 498 (1893).
[8] K. W. Charitschkoff, J. Russ. Chem. Soc. **39**, 230 (1907).

TABLE I. Products of photolysis using mercury arcs. Gases are in cc at N.T.P.

Expt. number	28	29	34	36	39	40	41	49	50	52
Light, arc and dist. in cm	II, 3	II, 3	Cd arc	I, 6	III,* 8	I,* 8	I, 8	IV, 8	IV, 8	IV, 8
Initial $(CH_3NH)_2$	4.85	5.23	4.00	4.40	4.55	1.10	3.92	0.95	1.70	0.35
Time (min.)	140	22	70	30	53	325	25	60	61	60
Press. increase (mm)	20.6	8.5	15.0	14.0	6.2	2.7	3.2	2.0	3.1	0.7
H_2	1.74	0.72	1.32	1.05	0.45	0.22	0.27	0.17	0.32	0.10
CH_4	1.59	0.50	1.02	1.00	0.32	0.15	0.16	0.12	0.22	0.06
C_2H_6	0.20	0.05	0.10	0.08	0	0	0	0	0	0
N_2	1.92	0.49	1.32	1.02	0.35	0.22	0.16	0.17	0.32	0.11
NH_3	0.55	0.46	0.38	0.50	0.30	0.16	0.30	0.05	0.10	0.05
CH_3NH_2	0.45	0.40	0.35	0.37	0.25	0.01	0.27	0	0	0
$(CH_3NH)_2$ used	2.80	1.46	2.50	2.00	1.10	0.49	0.75	0.22	0.50	0.20
% decomposed	58.0	28.0	62.0	45.5	23.2	44.5	19.1	23.2	29.4	57.0

* Crystal quartz lens used.

TABLE II. Dimethylhydrazine quantum yields using Arc IV. Gases are in cc at N.T.P.

Expt. number	46	47	48	48A	51	53	53A	54	55	56
Initial $(CH_3NH)_2$	4.17	4.03	3.90	3.50	3.20	4.03	2.50	4.13	4.02	4.03
Time (min.)	60	60	60	60	60	10	10	90	5	60
Press. increase (mm)	5.0	8.4	7.1	6.4	6.6	1.2	1.5	14.0	0.5	6.3
H_2	0.47	0.92	0.74	0.60	0.60	0.17	0.14	0.95	0.10	0.52
CH_4	0.25	0.58	0.46	0.39	0.38	0.07	0.05	0.74	0.04	0.32
N_2	0.24	0.57	0.46	0.40	0.42	0.12	0.11	0.78	0.08	0.31
NH_3	0.30	0.49	0.41	0.40	0.35	0.10	0.12	0.37	0.08	0.32
CH_3NH_2	0.25	0.37	0.35	0.35	0.25	0.08	0.10	0.20	0.04	0.26
$(CH_3NH)_2$ used	1.15	1.70	1.40	1.20	1.22	0.30	0.25	1.87	0.14	1.10
ϕ on $(CH_3NH)_2$	0.28	0.32	0.33	0.29	0.29	0.42	0.35	0.30	0.40	0.26
ϕ on H	0.23	0.34	0.35	0.29	0.29	0.47	0.39	0.30	0.60	0.26
% decomposed	27.6	41.2	36.0	34.3	38.1	7.4	10.0	45.2	3.5	27.3

TABLE III. Products of photolysis using spark sources.

Expt. number	33	33A	42	43	44	44A	45	57	58	58A	59	65	66	67
Light	Full Zn	Full Zn	Full Zn	Zn F.Q.	Full Cd	Full Cd	Full Zn	Zn F.Q.	Full Al	Full Al	Zn F.I.	Zn F.Q.	Zn F.I.	Full Al
Init. $(CH_3NH)_2$	4.40	4.35	4.35	3.90	3.86	4.25	4.40	3.49	2.96	3.50	3.68	3.70	3.80	3.90
Time (min.)	105	105	161	450	143	140	150	310	194	200	300	340	780	60
Press. increase (mm)	8.3	8.4	1.5	0	8.7	8.8	7.5	5.9	7.3	6.8	3.3	3.0	3.5	2.0
H_2	0.54	0.56	0.15	0.04	0.58	0.60	0.51	0.22	0.54	0.52	0.11	0.24	0.20	0.22
CH_4	0.35	0.35	0.05	0.01	0.34	0.35	0.30	0.12	0.33	0.32	0.05	0.12	0.10	0.10
N_2	0.33	0.36	0.09	0.02	0.38	0.36	0.36	0.20	0.36	0.33	0.07	0.20	0.15	0.15
NH_3	0.42	0.47	0.15	0.04	0.45	0.43	0.41	0.27	0.27	0.30	0.18	0.25	0.10	0.17
CH_3NH_2	0.40	0.44	0.10	0	0.40	0.40	0.41	0.25	0.21	0.23	0.06	0.18	0.10	0.10
$(CH_3NH)_2$ used	1.20	1.20	0.35	0.10	1.25	1.12	1.09	0.50	1.00	1.02	0.24	0.60	0.48	0.40

F.I., focal isolation; F.Q., fused quartz lens (>2400A); gases are in cc at N.T.P.

match the precipitate from a run. Equal amounts of dimethyl and methylamines and ammonia gave a lighter color and only with traces of dimethylamine could a color match be obtained. Dimethylhydrazine gives a brown precipitate with Nessler's reagent which is easily distinguishable.

The presence of a primary amine was proved by the pink color of the Rimini test using acetone and sodium nitroprusside. The absence of a secondary amine was proved by the failure of the Simon test using acetaldehyde and sodium nitroprusside.

The separation finally adopted was based on the François[9] observation that ammonia forms a compound with yellow mercuric oxide while methylamine does not. Tests on mixtures confirmed the observation. The mercuric oxide should be kept evacuated since exposure to mercury vapor was found to reduce its effectiveness. This method of analysis was checked

[9] M. François, Comptes rendus **144**, 857 (1907).

against that used by Emeleus and Jolley[10] of distilling the mixture of methylamine and ammonia in vacuum from −118°C to liquid nitrogen temperature, the ammonia being volatile. The two methods agreed reasonably closely. Combustion of the material not absorbed by mercuric oxide showed it to be methylamine. All tests for the presence of cyanide, possibly as hydrogen cyanide, failed.

During photolysis small amounts of a brown liquid formed on the sides of the reaction cell. After removal of the gases after an exposure, this liquid was found to be soluble in water, the solution being alkaline to methyl red. The solution reduced Fehling's solution in the cold while dimethylhydrazine only reduced it on warming. Methyl methyleneimine would hydrolyze to formaldehyde and thus reduce Fehling's solution in the cold.

Table I presents a summary of the experiments using various mercury arcs under varying conditions.

Quantum Yield

Since with Arc IV ninety-eight percent of the ultraviolet radiation is concentrated in the 2537A resonance line and the fused quartz lens eliminated the 1850A line, the amount of radiation at 2967, 3126, and 3132A which could be absorbed by uranyl oxalate must be extremely small. For complete absorption, as at 2537A, by both uranyl oxalate and dimethylhydrazine, the quantum yield of dimethylhydrazine is given by the ratio of the moles of the hydrazine decomposed to I_0, the average number of quanta per sec. entering the actinometer solution; this is given by the relation $I_0 = NM/\phi t$ where N is the Avogadro number, M is the number of moles of uranyl oxalate decomposed in t sec., and ϕ is the quantum yield, given by Brackett and Forbes[11] as 0.63 ± 0.03.

Quantum yields were also calculated on the basis of atomic hydrogen by considering each hydrogen molecule found by analysis as representing two hydrogen atoms from two dimethylhydrazine molecules. Table II presents a summary of the data obtained.

[10] H. J. Emeleus and L. J. Jolley, J. Chem. Soc. p. 1612 (1935).

[11] F. P. Brackett and G. S. Forbes, J. Am. Chem. Soc. 55, 4459 (1933).

It is apparent that the over-all quantum yield at 2537A is 0.3.

From the data in Tables I and II it may be seen that the amount of dimethylhydrazine decomposed being a linear function of the time at a given initial pressure and intensity of illumination is proportional to the total light absorbed at least up to 40 percent decomposition. In Fig. 2 data from Tables I and II are presented graphically as the volume of product formed per unit amount dimethylhydrazine decomposed against the percentage decomposition. The small volumes of products and of dimethylhydrazine decomposed for runs at low pressures justify their exclusion from the graph. For the same reasons runs with 10 percent decomposition or less are not reliable. The hydrogen, methane, and nitrogen values are all high under these conditions suggesting that the error lies in the value of the hydrazine decomposed rather than in the gas analysis, since then the nitrogen, determined by difference, should be low for high methane and hydrogen values.

A material balance calculated for the more reliable runs in Tables I and II, of the differences between the hydrazine decomposed and the C, N,

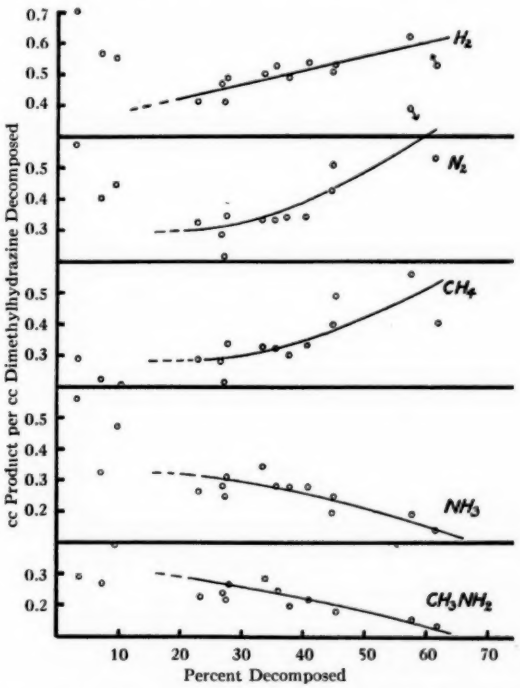


FIG. 2. Variation of products with amount decomposed.

and H found in the products, expressed as a C_2 compound since carbon appears in a minimum number of products, gave $C_2H_{5.1\pm0.4}N_{1.1\pm0.1}$. The lowest value for H in any run was 3.5 and the highest 5.7, while N varied between 0.6 and 1.4.

Effect of Wave-Length

The products of photolysis using the spark between electrodes of various metals are presented in Table III. Runs were made using the zinc spark with and without the fused quartz lens between cell and source to study the effect of the zinc spectrum above 2400A as compared with the full spectrum. No significant variations are observable. This was also true when the zinc spectrum between 2030 and 2140A, obtained by focal isolation, was used. The aluminum spectrum is rich in lines below 1900A, a region in which both ammonia and methylamine are known to decompose rapidly. This would account for the low values of these products found in runs with aluminum electrodes.

Effect of Propylene

Since propylene does not readily form active intermediaries or enter into a chain reaction it has been used to detect the presence of hydrogen atoms in the photolysis of hydrazine and amines. Bamford[12] found a large reduction in the amount of hydrogen and the production of propanes and isomeric hexanes. Several runs were made in presence of added propylene. The data are given in Table IV in which the values of propylene added, residual propylene, and propane have all been corrected for the propane originally present in the added propylene. It was not possible to determine the hexanes directly; the amount was estimated as half the difference between the propylene used and the propane formed. The fractionation of the product gases from these runs was made as before. The sec-butyl chloride fraction was treated with fuming sulfuric acid to remove propylene. Combustion of the residue gave a carbon value slightly less than that for propane, probably due to dissolved methane. The dry ice fraction was treated as before to remove ammonia and methylamine, and then subjected to combustion. The carbon value found was 3.

[12] C. H. Bamford, J. Chem. Soc. p. 17 (1939).

TABLE IV. Products of photolysis in presence of propylene. Gases are in cc at N.T.P.

Expt. number	37	38	60	61	63
Light	Hg Arc II	Hg Arc I	Full Zn	Zn F.Q.	Full Zn
Initial $(CH_3NH)_2$	4.15	4.77	3.24	3.70	3.89
Time (min.)	125	90	140	189	175
Press. increase (mm)	10.5	20.5	4.8	1.4	0.3
C_3H_6 added	10.50	5.30	7.85	10.90	17.70
C_3H_6 left	9.25	4.00	7.10	10.55	17.25
H_2	0.29	0.92	0.11	0.03	0.01
CH_4	0.78	1.51	0.32	0.12	0.14
C_3H_8	0.25	0.33	0.15	0.05	0.05
C_6H_{14} (estim.)	0.5	0.5	0.3	0.15	0.2
N_2	0.85	1.81	0.25	0.20	0.12
NH_3	0.50	0.56	0.23	0.20	0.16
CH_3NH_2	0.60	0.70	0.30	0.23	0.19
$(CH_3NH)_2$ used	1.9	2.9	0.8	0.5	0.5

With large amounts of propylene, a higher stationary concentration of propyl radicals is to be expected and a lower hydrogen atom concentration. The table shows larger amounts of hexane relative to propane at higher propylene concentrations. Comparison with the results obtained in the absence of propylene shows that hydrogen production is very considerably reduced and that the over-all decomposition is also retarded.

DISCUSSION

The absorption spectrum of ammonia[13] shows a region of double-headed diffuse bands beginning at 2400A and extending to 1665A. Below 1665A the bands are sharp and show rotational fine structure. True continuous absorption begins around 1200A and below 1150A becomes so intense as to obscure the bands. The spectrum is attributed to an excitation of a non-bonding electron associated with the nitrogen atom. The spectrum indicates the existence of a predissociation phenomenon over the range in which photochemical studies have been made, with the primary reaction postulated as

$$NH_3+h\nu\rightarrow NH_3^*\rightarrow NH_2+H.$$

In the alkylamines absorption begins at 2300–2500A and extends in a series of diffuse bands to shorter wave-lengths. The intensity and diffuseness increases as the wave-lengths decrease; with dimethyl and trimethylamine the bands eventually give rise to a continuum. It appears from the similarity of the spectra of the amines and

[13] A. B. F. Duncan, Phys. Rev. **47**, 822, 866 (1935); **49**, 211 (1936).

ammonia that the cause of both is the same and Bamford[12] has postulated the primary rupture to be

$$CH_3NH_2 \xrightarrow{h\nu} CH_3NH + H.$$

In hydrazine also, faint diffuse absorption bands, probably due to predissociation, begin at 2370A. Bamford,[14] from a study of hydrazine photolysis in presence of propylene and of nitric oxide, has furnished strong evidence that the primary rupture involves the scission of a hydrogen atom from the hydrazine molecule.

By analogy then, the continuum in dimethylhydrazine is probably "experimental" and involves a predissociation. Light absorption results in the formation of an excited molecule which might conceivably undergo one or more of the following dissociations

$$CH_3NHNHCH_3 \rightarrow \begin{cases} H + R & (a) \\ CH_3 + R' & (b) \\ 2CH_3NH. & (c) \end{cases} \quad (1)$$

Reaction (1a) would seem, by analogy, to be the most probable reaction and is supported by the reduction in hydrogen production when propylene is present, as also by the agreement between the quantum yields calculated directly and on the basis of H. Inspection of Fig. 2 shows that hydrogen production exceeds that of any other product. The presence of relatively large amounts of methane, as also of small amounts of ethane in the runs made with the hot arc for longer periods, indicates the probability of the presence of methyl radicals. The photolysis of methylamine does not give appreciable quantities of methane whence it is necessary to consider (1b) as a source of these radicals. Reaction (1c) is unlikely since it would be expected, initially at any rate, to yield larger amounts of methylamine than of the other products. A dissociation of an activated dimethylhydrazine molecule directly into stable molecules likewise seems improbable since larger amounts of methane relative to hydrogen than are observed would be expected.

Little can be said definitely concerning the fate of the atoms or radicals produced in reactions (1a) and (1b). Hydrogen and methane might result from such reactions as

$$H + CH_3NHNHCH_3 \rightarrow H_2 + R, \quad (2a)$$

$$CH_3 + CH_3NHNHCH_3 \rightarrow CH_4 + R. \quad (2b)$$

An alternative, though less likely reaction, for hydrogen production would be by recombination of hydrogen atoms. The relative weakness of the N−N bond as compared with the N−C and N−H bonds suggests that the reaction of hydrogen atoms and dimethylhydrazine might cause a rupture such as

$$H + CH_3NHNHCH_3 \rightarrow CH_3NH_2 + CH_3NH.$$

Henkin and Taylor[2] present evidence that the energy of activation of this reaction is about 8 kcal. The fate of the radical CH_3NH might be pictured as

$$2CH_3NH \rightarrow CH_3N:CH_2 + NH_3.$$

It is seen from Fig. 2 that while H_2, N_2, and CH_4 increase with increase in the percentage decomposition, NH_3 and CH_3NH_2 decrease. This decrease is to be expected since each undergoes photolysis at the wave-lengths used and gives hydrogen and nitrogen as the chief products. The increase in methane with increased percentage decomposition points to the accumulation in the system of some product which is a better potential source of methane than dimethylhydrazine itself. In the absence of these secondary decompositions the curves should be horizontal lines. In any case they must start out, at small decompositions with a horizontal tangent. It is not unreasonable from the figure that all the curves might originate at a value around 0.3; that is, stoichiometrically three molecules of dimethylhydrazine are involved in the production of one molecule each of the measured products. A material balance between three molecules of dimethylhydrazine and one molecule each of H_2, N_2, CH_4, NH_3, and CH_3NH_2 leaves a discrepancy of 4C, 10H, and 2N atoms. Methyl methylene imine $CH_3N:CH_2$ is known to polymerize to a dimer, its presence has been indicated among the products, and the actual material balance on runs in Tables I and II supports this possibility. If reactions of types (1) and (2) are correct, the significance of the three molecules of dimethylhydrazine involved must mean that the radicals R and R′ produced in

[14] C. H. Bamford, Trans. Faraday Soc. **35**, 568 (1939).

reactions (1) and (2) require a third molecule of the hydrazine for ultimate decomposition giving over-all

$$2R+CH_3NHNHCH_3=CH_4+N_2+NH_3+CH_3NH_2+(CH_3N:CH_2)_2,$$

or

$$2R'+CH_3NHNHCH_3=H_2+N_2+NH_3+CH_3NH_2+(CH_3N:CH_2)_2.$$

The observed quantum yield of 0.3 as opposed to the analytical yield of 3 is probably due to a reversal of (1a) or its equivalent. Henkin and Taylor[2] in the study of the addition of hydrogen atoms to azomethane found complete conversion of the azomethane to dimethylhydrazine even when the ratio of atomic hydrogen to azomethane was much less than 0.5 whereas, had the addition occurred entirely by steps involving atomic hydrogen, the ratio would have to have been at least two. It was obvious that some of the hydrogenation was occurring through hydrogen molecules and the reaction of the radical formed by addition of one hydrogen atom to azomethane, with hydrogen:

$$H_2+CH_3NHNCH_3\rightarrow CH_3NHNHCH_3+H,$$

appeared very plausible as a Walden inversion type reaction. This reaction is equivalent to a reversal of (1a) and, since the maximum energy of activation of addition of hydrogen to azomethane was found to be 3–4 kcal., will occur quite readily reducing the quantum yield as found.

The recent observation of Birse and Melville[15] that the efficiency of the primary process in ammonia photolysis is 0.58 rather than one indicates, since no fluorescence has been detected in ammonia and the pressure dependence of the quantum yield is less than would be expected for any collisional process, that there is some mechanism for internal degradation of the energy. This might also be present in the hydrazine molecule.

[15] E. A. B. Birse and H. W. Melville, Proc. Roy. Soc. **A175**, 164 (1940).

AUGUST, 1942 JOURNAL OF CHEMICAL PHYSICS VOLUME 10

The Diffraction of X-Rays by Liquid Oxygen*

P. C. SHARRAH AND N. S. GINGRICH
University of Missouri, Columbia, Missouri
(Received May 17, 1942)

The x-ray diffraction pattern of liquid oxygen at 89°K shows an intense peak at sin $\theta/\lambda=0.157$, and very weak peaks at 0.35 and at 0.5; that at 62°K shows an intense peak at 0.159, and weak peaks at 0.35 and at 0.5. The atomic distribution curves have peaks at 1.3A, 2.2A, 3.4A, and 4.2A for 89°K, and at 1.25A, 2.15A, 3.2A, and 4.1A for 62°K. The first peaks are due to the interatomic distance in the diatomic molecule, the second may be due to triatomic oxygen, and the others may be due to higher molecular aggregates.

INTRODUCTION

A STUDY of the diffraction of x-rays by liquid oxygen is of interest not only for its contribution to the field of x-ray diffraction, but also for the information it may supply to our knowledge of molecular aggregates in liquid oxygen. The present work was undertaken to obtain as complete a diffraction pattern of liquid oxygen as possible and to analyze this pattern for the atomic distribution curve. This has been done for liquid oxygen at 89°K and at 62°K.

One of the earliest studies in the field of x-ray diffraction by liquids was done with liquid oxygen by Keesom and de Smedt. In their first work,[1] copper radiation filtered through nickel was used in an evacuated camera of 2.75-cm radius. Two peaks in the intensity curve were reported, the main peak at sin $\theta/\lambda=0.151$ and a weak second-

* A preliminary report of this work was presented at the December, 1940 meeting of the American Physical Society.

[1] W. H. Keesom and J. de Smedt, Proc. Amst. Akad. Sci. **25**, 118 (1922).

ary peak at sin $\theta/\lambda = 0.254$. In their second work,[2] molybdenum radiation filtered through zirconium was used, enabling them to include larger values of sin θ/λ. They reported, in this case, the main peak at sin $\theta/\lambda = 0.154$, a weak second peak at sin $\theta/\lambda = 0.239$ on six out of nine films, and a somewhat questionable peak at sin $\theta/\lambda = 0.35$. Using Ehrenfest's relation,[3] $a = 7.72\lambda/4\pi \sin\theta$, and the angle θ corresponding to the main peak, they calculated "a" and correlated this with the distance between molecules in a close-packed system. The primary result of this work was that it indicated that the intense peak was probably due to interference of x-rays diffracted by neighboring molecules.

In view of the recent advances in both experimental technique and theoretical interpretation in this field, a repetition and extension of the early work of Keesom and de Smedt is clearly desirable. This attempt at an application of these newer techniques and interpretation has been made in the present work.

EXPERIMENTAL

Two types of system were used to obtain a sample of liquid oxygen, one which will be referred to as an open system, and one which will be referred to as a closed system. In using the open system, a large amount of liquid oxygen was prepared from tank oxygen with no further purification, and from pressure and temperature measurements, assuming it to have nitrogen as an impurity, its purity was found to be better than 96 percent. This liquid was transferred to a Dewar flask fitted at the bottom with a very thin-walled Pyrex glass tube 7.7-mm diameter in an evacuated camera of 9.53-cm radius, as shown in Fig. 1. As the oxygen evaporated, more was added, and for the low temperature work, a pump reduced the pressure over the oxygen. In using the closed system, which was nearly the same as that used for argon,[4] tank oxygen was given some further purification and introduced into the liquid air-cooled capillary at approximately one atmosphere. In this case, the reduced temperature was attained by pumping on the liquid air in the Dewar flask.

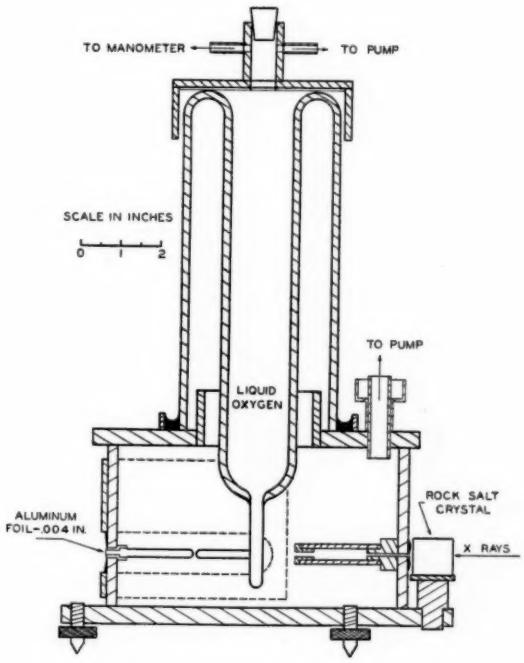


FIG. 1. A diagram of the evacuated camera using the "open system."

Mo $K\alpha$ radiation reflected from a rocksalt crystal was collimated in part by the crystal, and in part by two slits and the width of the beam at the sample was about 1.5 mm. A photographic film recorded the diffracted x-rays and exposures of 20 to 30 hours gave adequate densities on the films. A Leeds and Northrup recording microphotometer was used to convert film densities to relative intensities, in the usual manner.

The relative intensity curves were corrected for polarization[5] and for absorption. The absorption correction used was that for a cylindrical sample completely bathed in a parallel beam of x-rays, as discussed by Blake.[6] In the present case, the sample was only partially irradiated, and hence this correction is approximate. However, it has been shown[7] that for sodium, with cylindrical samples having diameters of from one to five times the width of the beam, there is little difference in the ratio of the intensity of scatter-

[2] W. H. Keesom and J. de Smedt, Proc. Amst. Akad. Sci. **26**, 112 (1923).
[3] P. Ehrenfest, Proc. Amst. Akad. Sci. **17**, 1184 (1915).
[4] A. Eisenstein and N. S. Gingrich, Phys. Rev. **58**, 307 (1940).
[5] S. Katzoff, J. Chem. Phys. **2**, 841 (1934).
[6] F. C. Blake, Rev. Mod. Phys. **5**, 180 (1933).
[7] G. A. Mitchell, unpublished work.

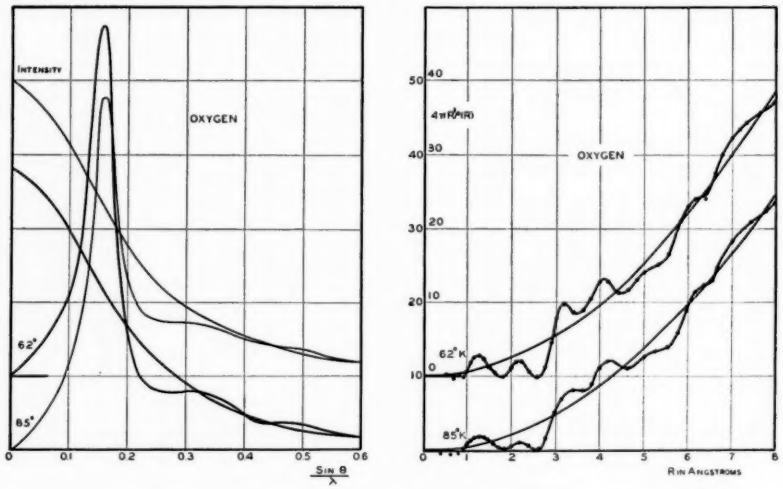


FIG. 2(a). (Left.) The x-ray diffraction patterns of liquid oxygen at 89°K and at 62°K. (b). (Right.) The atomic distribution curves for liquid oxygen at 89°K and at 62°K.

ing at the main peak to that at angles as large as were used here. It is believed, therefore, that this approximation has introduced no serious error. The curve corrected in this way and for incoherent radiation was fitted to the independent coherent scattering curve for making an analysis of the pattern.

Seven exposures were made with the oxygen at 89°K, and six were made at 62°K, this lower temperature being attained by pumping on the liquid oxygen or on the liquid air. Some of the films were heavily over-exposed in order to disclose features of the pattern at large angle where the intensity is normally very low. Typical curves are reproduced in Fig. 2(a). At each temperature, the pattern consists of one intense peak and two secondary prominences more in the nature of plateaus than of peaks. At 89°K, these occur at values of sin θ/λ of 0.157, 0.35, and 0.5, while in the 62°K case, they occur at 0.159, 0.35, and 0.5. The position of the main peak at 89°K, 0.157, is slightly higher than the two values given by Keesom and de Smedt, 0.151[1] and 0.154,[2] but it is believed that with the larger camera here used, the present value is likely to be more nearly correct. Their second peak at 0.254 or 0.239 could not be found and it is believed that there is no peak near this position. Their third peak[2] at 0.35, however, does check well with a prominence found in the present work.

ANALYSIS OF PATTERNS

Fourier analyses of the patterns were made according to the method of Debye and Menke[8] giving the atomic distribution of atoms about any atom in the liquid. The expression used is derived elsewhere[9] and is

$$4\pi R^2\rho(R)=4\pi R^2\rho_0+2R/\pi\int_0^\infty Si(S)\sin Rsds,$$

where $i(s)$ is a function obtained from the experimental data, $S=4\pi\sin\theta/\lambda$, θ is half the scattering angle, ρ_0 is the average density of atoms, $\rho(R)$ is the actual density of atoms, and R is the distance in angstroms from the origin atom. Four analyses were made for 89°K oxygen and five for 62°K oxygen. These various analyses were made for different patterns and for different fittings on a given pattern. In about half the cases, the integral was evaluated by plotting $Si(S)\sin Rs$ against S and integrating graphically, and for the remaining ones, the integral expression was evaluated by the method of trigonometric interpolation.[10] As is to be expected, the two methods gave very closely the same result for a given curve and fitting. The large number of analyses was made because of the unusual appearance of

[8] P. Debye and H. Menke, Physik. Zeits. **31**, 797 (1930).
[9] B. E. Warren, J. App. Phys. **8**, 645 (1937).
[10] C. Lanczos and G. C. Danielson, Phys. Rev. **55**, 242 (1939).

the distribution curve with many prominent peaks. An averaged curve for each temperature is shown in Fig. 2(b). In all the curves obtained from these analyses, most of the essential features shown in the figure were reproducible up to a little over 5A, but beyond that distance the various analyses differed considerably, and hence, no confidence is placed in the curves beyond 5A. At less than this distance, the first four peaks appeared in all cases at approximately the same positions but with somewhat different heights. This feature appeared to be unusually dependent upon the fitting of the corrected experimental curve to the independent coherent curve for the peaks between 3A and 5A. The first peak was reproducible to within about 5 percent of the reported value for position and to within about 10 percent of the reported value for the area and the weak second peak persisted under all fittings with variations in position and area slightly greater than for the first peak.

DISCUSSION OF RESULTS

The positions of the peaks for the 89°K case are 1.3A, 2.2A, 3.4A, and 4.2A, and for the 62°K case they are 1.25A, 2.15A, 3.2A, and 4.1A The area under the first peak, for 89°K, is 1.08, and that for 62°K is 1.18, indicating that there is one nearest neighbor, or that liquid oxygen has a fundamental diatomic aggregate closely similar to the diatomic molecule in the gas, for which each atom has one nearest neighbor at about 1.22A.[11,12] From this analysis, one cannot be sure that there is a significant difference between 89°K and 62°K, in the positions of this first peak, i.e., between 1.3A and 1.25A. The area under these peaks is slightly more than one, but it is believed that the precision and reproducibility of this peak is hardly great enough to give this difference much significance. Liquid oxygen is the only element thus far to show a second discrete peak, which occurs here at 2.2A and at 2.15A in the cases of 89°K and 62°K, respectively, and the interpretation of this second peak, if it is real, is somewhat difficult. It is true that from the assigned structure of ozone,[13] O_3, the first interatomic distance is between 1.25A and 1.30A and that the second distance is about 2.2A, but the amount of O_3 present in liquid oxygen is normally very small. On the other hand, irradiation of liquid oxygen is known[14] to enhance the amount of O_3 present, which would contribute both to the presence of this peak and to an area under the first peak greater than one, since one atom of O_3 would have two neighbors at about 1.3A. Thus the distribution curves here obtained are consistent with the assumption that there was O_3 in the samples used, though the data are not considered sufficiently well known to warrant making quantitative deductions from this result.

The third and fourth peaks appear to be related to distances between oxygen atoms in larger aggregates of oxygen molecules. From studies of the magnetic susceptibility and of the absorption spectra of oxygen,[15] it appears probable that O_4 may exist in the liquid in a considerable amount. Vegard[16] has determined the crystal structure of the γ-form of crystalline oxygen, and he states that the x-ray diffraction pattern was typical of that for a cubic lattice with rotating molecules. The O_2 molecules are grouped into pairs along the trigonal axis and the distance between the two molecules of a pair is given as 3.48A. In the present work, two distances, about 3.4A and 4.2A, appear fairly predominantly, and it is not unlikely that the former may have some relation to Vegard's value, 3.48A, though an unambiguous interpretation of both these peaks appears somewhat difficult.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge financial support of this work by the American Academy of Arts and Science, the American Association for the Advancement of Science, and the Elizabeth Thompson Science Fund.

[11] A. H. Compton and S. K. Allison, *X-Rays in Theory and Experiment* (D. Van Nostrand Company, 1935), p. 161.
[12] G. Herzberg, *Molecular Spectra and Molecular Structure* (Prentice Hall, 1939), p. 491.
[13] W. G. Penney and G. B. B. M. Sutherland, Proc. Roy. Soc. **A156**, 672 (1936).
[14] J. W. Mellor, *A Comprehensive Treatise on Inorganic and Theoretical Chemistry* (Longmans, Green and Company, 1922), p. 880.
[15] M. and B. Ruheman, *Low Temperature Physics* (Cambridge University Press, 1937), p. 228.
[16] L. Vegard, Nature **136**, 720 (1935).

On the Absorption Spectrum of Some Polymethine Dyes

KARL F. HERZFELD
Department of Physics, Catholic University of America, Washington, D. C.
(Received April 29, 1942)

The energy levels of certain polymethine dyes are calculated approximately. The dyes consist of a chain of conjugated double bonds, with end groups which might or might not differ from each other. Ions and neutral molecules are treated. The calculations are performed both with the valence bond method (resonance between different structures) and the LCAO method, but without antisymmetrization and without taking electronic interaction into account. The effect of chain length and of the nature of the end groups is discussed.

I. INTRODUCTION

IN a series of papers, Brooker and Sklar[1] discuss the color, or in general, the first electronic transition in a series of dyes, mainly of cyanines.

The present paper has as its purpose an approximate quantum theoretical discussion, which, while insufficient to permit an actual calculation of the wave-length of the absorption, would at least permit us to understand the dependence on the chain length and on the nature of the end members of the molecules.

As is well known, there are essentially two methods to treat the electron states, the Heitler-London-Pauling-Slater method (abbreviated HLPS) and the molecular orbital method (Hückel-Hund-Mulliken, abbreviated LCAO). Beforehand, we cannot know which will fit better. It turns out, however, that the first part of the formal calculation is the same in both; only the numerical value of the constants involved and the final application of the formula are different.

Let us, as an example of the structures discussed in the HLPS method, consider the ion

$$\underset{H}{H_2\overset{+}{N}{=}C}-\underset{H}{\underset{|}{(C}}=\underset{H}{C)_n}-NH_2. \qquad (1)$$

It contains $N=2n+1$ CH groups. According to the HLPS viewpoint, there are two ground structures, namely besides (1)

$$H_2N-\underset{H}{\underset{|}{(C}}=\underset{H}{C)_n}-\underset{H}{C}=\overset{+}{N}H_2, \qquad (1')$$

the (negative) energy of which we call e. Besides the two ground structures there are $N+1$ higher structures

$$H_2N-\underset{H}{\overset{+}{C}}-\underset{H}{\underset{|}{(C}}=\underset{H}{C)}-NH_2, \qquad (2)$$

$$H_2N-\underset{H}{\overset{\vdots}{C}}-\underset{H}{\overset{+}{C}}-\underset{H}{\overset{\vdots}{C}}-\underset{H}{\underset{|}{(C}}=\underset{H}{C)_{n-1}}-NH_2. \qquad (3)$$

Formulas (2) and (3) do not have exactly the same energy, but in first approximation we neglect this difference and ascribe to the excited structures the energy f, which is smaller than e (i.e., the excited structures lie above the ground structures). Any structures lying still higher are neglected. This would not be justified for a numerical calculation, but, we hope, does not matter for the qualitative discussion attempted here.

Furthermore, it is assumed that only neighboring structures interact, (1) with (2), (2) with (3)$\cdots(N+1)$ with $N+2$. The interaction energy is called α and is assumed to be the same between any two states.

When the energy levels—of which there are $N+2$—are calculated, the HLPS method takes the lowest one as the ground state and the next higher one as the first excited state. The first absorption band is therefore given by the energy difference of the two lowest states. On the other hand, the molecular orbital method starts from the skeleton

$$\underset{H}{H_2\overset{+}{N}-\overset{+}{C}}-\underset{H}{\overset{+}{C}}\ldots\underset{H}{\overset{+}{C}}-\overset{+}{N}H_2 \qquad (4)$$

into which $N+1$ electrons are put. There will be $(N+2)$ states, of which $(N+1)/2$ are filled. The first transition takes an electron from state $(N+1)/2$ and puts it into $(N+3)/2$. However, the constants have entirely different meaning and

[1] For the literature, see the following paper by A. L. Sklar.

value in the HLPS and the LCAO method, even if they appear formally in the same manner in the equations.

The LCAO method consists in forming a molecular orbital of the whole molecule out of a sum of atomic orbitals for the different ions, $\overset{+}{\underset{H}{C}}$ or $\overset{+}{N}H$. Mathematically, this task is identical with that of the HLPS method.

II. SYMMETRICAL DYNES, $N>1$

For dyes like the ion (1), which have equal ends, the energy is determined by the determinant of order $N+2$

$$\begin{vmatrix} e-E & \alpha & 0 & 0 & 0 & 0 \\ \alpha & f-E & \alpha & 0 & 0 & 0 \\ 0 & \alpha & f-E & \alpha & 0 & 0 \\ 0 & 0 & & & & \\ & & & & \alpha & 0 \\ & & & \alpha & f-E & \alpha \\ & & & 0 & \alpha & e-E \end{vmatrix} = 0. \quad (5)$$

The problem of the dienes $e=f$ has been extensively treated.[2] There, however, the difference between the end members and the chain members is neglected, while the influence which this difference exerts is one of our principal interests. On the other hand, Lennard-Jones, Turkevich, and Coulson have taken into account the fact that the chain double bonds might alternatively have different lengths, a fact we neglect.

To solve (5) we write the classical differential equations of a system of coupled oscillators for which (5) is the secular equation with ω^2 instead of E. Calling the vibration coordinate of the end members X and Y and that of the sth chain members x_s, we have

$$\ddot{X} = -eX - \alpha x_1, \quad (6)$$

$$\ddot{x}_1 = -fx_1 - \alpha X - \alpha x_2, \quad (7)$$

$$\ddot{x}_s = -fx_s - \alpha(x_{s-1} + x_{s-1}), \quad s = 2, \cdots N-1 \quad (8)$$

$$\ddot{x}_n = -fx_n - \alpha Y - \alpha x_{n-1}, \quad (9)$$

$$\ddot{Y} = -eY - \alpha x_n. \quad (10)$$

This can be interpreted as the equation of motion of a vibrating chain in which the end members are tied to positions of equilibrium in space by quasi-elastic forces $-(e-\alpha)X$ or $(e-\alpha)Y$, the chain members with quasi-elastic forces $-(f-2\alpha)x_s$. In addition there act between neighboring particles destabilizing quasi-elastic forces $+\alpha(x_s-x_{s-1})$, $\alpha(X-x)$, etc.

It turns out that the system of differential Eqs. (6)–(10) is completely equivalent to one treated by Bartholomé and Teller.[3] These authors treat the vibrations of a carbon chain (quasi-elastic force constant k_c, mass m_c) with valence angle β, to which end members (quasi-elastic force constant k_E, mass m_E) are attached. (6)–(10) go over into their equations if one identifies as follows:

$$E:\omega^2; \quad e:k_E\left(\frac{1}{m_E}+\frac{1}{m_c}\right); \quad f:\frac{2k_e}{m_c}; \quad \alpha=\frac{k_c}{m_c}\cos\beta.$$

Furthermore, nowhere do they make use of the fact that $\cos\beta \leqq 1$, so that no limitation is put on the ratio of α/f. Bartholomé and Teller have not actually written the equations for the energy levels or discussed them. That shall be done here.

One finds as a general result, which can easily be confirmed from the following, that there are $N+2$ solutions, that $(N+3)/2$ are symmetric and $(N+1)/2$ antisymmetric, that the symmetric and antisymmetric levels alternate, with a symmetric one at the bottom and at the top. A symmetric solution is one for which

$$X = Y, \quad x_{[(N+1)/2]-s} = x_{[(N+1)/2]+s} \quad (11)$$

$$s = 1 \cdots \frac{N-1}{2};$$

an antisymmetric one has a minus sign on the left of both Eqs. (11). It follows that for all cases in the HLPS method and for an odd number of double bonds in the LCAO method, the first transition takes an electron from a symmetric state and puts it into an antisymmetric one.[4]

We introduce the quantity

$$R = \alpha/(e-f), \quad (12)$$

[2] E. Hückel, Zeits. f. Physik **70**, 204 (1931); **76**, 628 (1932); Zeits. f. Electrochemie **43**, 752 (1937). R. S. Mulliken, J. Chem. Phys. **7**, 364, 570 (1939). J. E. Lennard-Jones, Proc. Roy. Soc. **158**, 280 (1937). J. E. Lennard-Jones and J. Turkevich, Proc. Roy. Soc. **158**, 297 (1937). C. A. Coulson, Proc. Roy. Soc. **164**, 383 (1938); **169**, 413 (1939).

[3] E. Bartholomé and E. Teller, Zeits. f. physik. Chemie **B19**, 3 (1931).

[4] For an even number of double bonds, the highest filled one-electron state is odd, and the first excited state even. For the state of the ion as a whole, however, the statement in the text is true whatever the number of double bonds.

which gives the ratio of the interaction energy to the energy difference between the ground structures and the excited structures. Depending on the question whether $R>1$, $=1$, <1, the discussion will be different.

To solve (6)–(10), one writes ($A\sim e^{i\omega t}$) for symmetric states

$$x_s=A\cos\left(\frac{N+1}{2}-s\right)k, \tag{13}$$

for antisymmetric ones

$$x_s=A\sin\left(\frac{N+1}{2}-s\right)k. \tag{14}$$

k is the "wave vector" of the standing wave, and corresponds to π/λ in Bartholomé and Teller's paper. (In addition, they count s from the middle instead of from one end.)

Both (13) and (14) inserted in (8) give

$$E-f=2\alpha\cos k. \tag{15}$$

X is then found, from (6) as

$$X=x_1\frac{\alpha}{E-e}. \tag{16}$$

One gets, by inserting (15), (16), and (13) or (14) in (7), for the symmetric states

$$\frac{\cos\frac{N+1}{2}k}{\cos\frac{N+3}{2}k}=R, \tag{17}$$

for the antisymmetric ones

$$\frac{\sin\frac{N+1}{2}k}{\sin\frac{N+3}{2}k}=R. \tag{18}$$

If $R\geqq 1$, these equations give the appropriate number of solutions, but if $R\ll 1$, the number is too small by two. The additional solutions are found by putting[5]

$$x_s=A\{v^{[(N+1)/2]-s}\pm v^{-[(N+1)/2]+s}\}, \tag{13'}$$

where the upper sign is for the symmetric, the lower for the antisymmetric solution. (15) is then replaced by

$$E-f=\alpha\left(v+\frac{1}{v}\right). \tag{15'}$$

(15′) has, for a given E, two solutions v which are reciprocal. We decide to call v the one which is <1.

Then, (17) and (18) are replaced by

$$1-\frac{v}{R}=\pm v^{N+1}\frac{1-v^2}{1\pm v^{N+1}}. \tag{17'}$$

One can easily prove again that the symmetric solution is the lower one and that both solutions of (17′) are below[6] those of (17) and (18).

However, for R near, but smaller than 1, (17′) does not give an antisymmetric solution. In fact one finds, using the negative sign and putting $v=1$ (which is the largest value v can have),

$$R=1-\frac{2}{N+3}.$$

For R smaller than this value, (17′) provides both a symmetric and an antisymmetric solution. For

$$1-\frac{2}{N+3}<R<1.$$

(17′) provides the symmetric solution, while an additional antisymmetric solution comes from (18), with

$$\frac{N+1}{2}k<\frac{\pi}{2}.$$

Th. Foerster,[7] in a very interesting paper, has given the formula (13), (14) (15), (17), (18). He evaluates them for $R=0$, 1, ∞, which can be done exactly, and discusses the trend of the absorption frequencies from his results. He does not attempt to give approximate expressions.

The rest of the discussion will be made for $R\ll 1$, $R=1$, $R\gg 1$.

1. $R\ll 1$ or $R<1$, $N\gg 1$ or both conditions.

In the HLPS method, we are interested only in the lowest two solutions, i.e., the two given by

[5] This can be expressed, by writing $v=e^{-k}$, as

$$x_s=A\cosh\left(\frac{N+1}{2}-s\right)k,\quad x_s=A\sinh\left(\frac{N+1}{2}-s\right)k.$$

(15′) is then $E-f=2\alpha\cosh k$.

[6] $v+1/v$ is larger than 2, except for $v=1$; $2\cos k$ is smaller than 2, except for $k=0$.

[7] Th. Foerster, Zeits. f. physik. Chemie **B47**, 245 (1940).

(17′). Then, in zero approximation, the right side is zero and

$$1-\frac{v}{R}=0.$$

For the next approximation, this is inserted in the right and v^{N+1} in the denominator neglected. Then[8]

$$1-\frac{v}{R}=\pm(1-R^2)R^{N+1},$$

$$v=R\mp(1-R^2)R^{N+2}, \tag{19}$$

$$E-f=\alpha\left(\frac{1}{R}+R\right)\pm\frac{\alpha}{R}(1-R^2)^2R^{N+1}$$

and finally

$$E=e+\frac{\alpha^2}{e-f}\pm(e-f)(1-R^2)^2R^{N+1}, \tag{20}$$

so that we see: the center of the two lowest levels is depressed below e by the amount $\alpha^2/(e-f)$, which is independent of N, while the splitting depends on N. The energy difference between these two levels, which in the HLPS view gives the first transition, is

$$\Delta E=2\alpha(1-R^2)^2R^N. \tag{21}$$

(21) says that the wave-length of the absorption band goes to infinity exponentially with N.

Our results have a close connection with the assumption of Pauling[9] that the large number of intermediate structures can be replaced by an increased overlap. That is so according to our calculations, but the quantitative side needs some discussion. In the present case, the dependence on N is of the same form which it would have in the case of an increased overlap integral. The latter usually has the form $C\exp(-r/\rho)$ where C and ρ are constants and r the distance between the overlapping atomic orbitals. If we now write for the chain $r=r_1+Nr_2$, the overlap integral would be $(C\exp-r_1/\rho)(\exp-r_2/\rho)^N$ which is of the form (21) with R instead of $(\exp-r_2/\rho)$. However, Pauling is not interested in the dependence on chain lengths, but in the influence of different chain endings, and therefore he expects the variation with e to be proportional to $(e-f)^{-1}$, i.e., expects a transition frequency $\sim R$ for fixed N and α. None of the formulas discussed in this chapter behaves quantitatively in this manner; only those found with the assumptions of Chapter V do.

In the HHM view, the lowest transition is the one which takes an electron from the topmost full band, the $(N+1)/2$th one, and puts it into the lowest empty one, i.e., the $(N+3)/2$th one.

We have therefore to investigate the solutions of (17) and (18). If R is very small, the numerators of (17) and (18) will be very small; therefore

$$\frac{N+1}{2}k=r\frac{\pi}{2}+\delta, \quad (r=1\cdots N) \tag{22}$$

r odd for symmetric solutions (i.e., (17)), even for antisymmetric ones.

Then (17) and (18) can be written in the same form

$$\sin\delta=R\sin\left(\frac{N+3}{N+1}\delta+\frac{r\pi}{N+1}\right). \tag{23}$$

For the highest filled state, $r=(N-1)/2$ and

$$E=f+2\alpha\cos\left(\frac{N-1}{N+1}\frac{\pi}{2}+\frac{2\delta_1}{N+1}\right)$$

$$=f+2\alpha\sin\frac{\pi-2\delta_1}{N+1}, \tag{24}$$

$$\sin\delta_1=R\sin\left(\frac{N-1}{N+1}\frac{\pi}{2}+\frac{N+3}{N+1}\delta_1\right)$$

$$=+R\cos\left(\frac{\pi}{N+1}-\frac{N+3}{N+1}\delta_1\right). \tag{24′}$$

For the first empty state, $r=(N+1)/2$ and

$$E=f-2\alpha\sin\frac{2\delta_2}{N+1}, \tag{25}$$

$$\sin\delta_2=R\cos\frac{N+3}{N+1}\delta_2. \tag{25′}$$

Up to now, the formulae are exact.

[8] This is a good approximation provided R^{N+1} is small compared to one. E.g., if $R=\frac{1}{2}$ and $N=3$ (a single C=C group), $R^{N+1}=\frac{1}{16}$ and the zero approximation $v=R$ is only wrong by 5 percent. The split, of course, is wrong by a greater error (about 20 percent). For $N=5$, the error in the split is less than 7 percent. It would be easy to improve the calculations for small N but not worthwhile, because of the simplified assumptions.

[9] L. Pauling, Proc. Nat. Acad. Sci. **25**, 577 (1939).

If R is very small and N arbitrary, (24′) and (25′) give approximately

$$\delta_1 = R\cos\frac{\pi}{N+1}, \quad \delta_2 = R,$$

$$\Delta E = 2\alpha\left\{\sin\frac{\pi}{N+1} - \frac{2R}{N+1} \times\cos^2\frac{\pi}{N+1} + \frac{2R}{N+1}\right\}$$

$$= 2\alpha\sin\frac{\pi}{N+1} + \frac{2R}{N+1}\sin^2\frac{\pi}{N+1}. \quad (26)$$

If N is very large and R has any value <1, (24′) and (25′) give approximately

$$\text{tg}\,\delta_1 = R\cos\left(\frac{\pi}{N+1}\right)\left(1+R\sin\frac{\pi}{N+1}\right)$$

$$\text{tg}\,\delta_2 = R.$$

The energy difference is, in first approximation,

$$\Delta E = 2\alpha\frac{\pi}{N+1}\left(1+2\frac{\delta_2-\delta_1}{\pi}\right)$$

or, if R is small (27)

$$\Delta E = 2\alpha\frac{\pi}{N+1}\left(1+2\frac{R^2}{N+1}\right).$$

In both (26) and (27), the wave-length goes to infinity[10] with increasing N. The principal term, which is independent of the nature of the ends, goes with the first negative power of $N+1$.

2. For $R=1$, (17) and (18) can be solved exactly, by

$$\frac{N+1}{2}k = r\pi - \frac{N+3}{2}k$$

or (28)[11]

$$k = \frac{r\pi}{N+2}, \quad r = 0, 1\cdots N+1$$

where r even is valid for the symmetric, r odd valid for the antisymmetric states.

The two lowest states are

$$E = f+2\alpha = e+\alpha \text{ (symmetric)} \quad (29)$$

and

$$E = f+2\alpha\cos\frac{\pi}{N+2} = e + \alpha\left(2\cos\frac{\pi}{N+2}-1\right) \text{ (antisymmetric).} \quad (29')$$

The energy difference for the first transition according to HLSP is

$$\Delta E = 2\alpha\left(1-\cos\frac{\pi}{N+2}\right) = 4\alpha\sin^2\frac{\pi}{2(N+2)}. \quad (30)$$

For very large N, the wave-length of the absorption band goes to infinity like N^2 [(more precisely, it is $hc/4\alpha\sin^{-2}\pi/2(N+2)$]. For the LCAO method the lowest ΔE is given by the difference between the $(N+1)/2$th level and the $(N+3)/2$th one, i.e., between the levels with $r=(N-1)/2$ and $(N+1)/2$; therefore

$$\Delta E = -2\alpha\left(\cos\frac{(N+1)\pi}{2(N+2)} - \cos\frac{(N-1)\pi}{2(N+2)}\right)$$

$$= 4\alpha\cos\frac{\pi}{(N+2)}\sin\frac{\pi}{2(N+2)}. \quad (31)$$

The wave-length of the absorption band goes to infinity with increasing N, but with the first power.

3. For $R>1$, (17) and (18) provide all the energy levels, i.e., (13) and (14) all the solutions. It is now better to take the reciprocals of (17) and (18).

$$\frac{\cos\frac{N+3}{2}k}{\cos\frac{N+1}{2}k} = \frac{1}{R} \text{ (symmetric),} \quad (17'')$$

$$\frac{\sin\frac{N+3}{2}k}{\sin\frac{N+1}{2}k} = \frac{1}{R} \text{ (antisymmetric).} \quad (18')$$

[10] The expression for ΔE is really a power series in $(N+1)^{-n}$. If the wave-length is calculated, the first term is proportional to $N+1$, the next term is independent of N. The first two terms of the expressions for the wave-length, taken together, are therefore proportional to $N+C$, where C is a constant independent of N. The following term is $\sim(N+1)^{-1}$.

[11] $r=N+2$ is odd and $k=\pi$ is not a solution of (18).

We expect now

$$\frac{N+3}{2}k=r\frac{\pi}{2}+\delta \tag{32}$$

with odd r for the symmetric, even r for the antisymmetric states.

Then (17″) and (18′) are again expressed in the single formula

$$\sin\delta=-\frac{1}{R}\sin\left(\frac{r\pi}{N+3}-\frac{N+1}{N+3}\delta\right). \tag{33}$$

$r=0$ and $r=N+3$ give no solution $\delta<\pi/2$ which (32) presupposes. r goes therefore from 1 to $N+2$, the two lowest states being $r=1$ and $r=2$ (HLPS) while for HHM the highest occupied state is $r=(N+1)/2$, the lowest unoccupied state $r=(N+3)/2$. We approximate again for $R\gg1$. (33) is then

$$\delta=-\frac{1}{R}\sin\left\{\frac{r\pi}{N+3}+\frac{N+1}{N+3}\frac{1}{R}\sin\frac{r\pi}{N+3}\right\} \tag{34}$$

and the energy

$$E=f+2\alpha\cos\left\{\frac{\pi r}{N+3}+\frac{2\delta}{N+3}\right\}$$

$$=f+2\alpha\cos\frac{r\pi}{N+3}\cdot\left[1-\frac{2}{R^2(N+3)^2}\cdot\sin^2\frac{r\pi}{N+3}\right]$$

$$+\frac{4\alpha}{R(N+3)}\cdot\sin\frac{r\pi}{N+3}$$

$$\cdot\sin\left\{\frac{r\pi}{N+3}+\frac{N+1}{N+3}\frac{1}{R}\sin\frac{r\pi}{N+3}\right\}.$$

Or

$$E=f+2\alpha\cos\frac{r\pi}{N+3}+4\frac{e-f}{N+3}\left(\sin^2\frac{\pi r}{N+3}\right)$$

$$\times\left[1+\frac{1}{R}\frac{N}{N+3}\cos\frac{r\pi}{N+1}\right]. \tag{35}$$

The first two terms would appear in the pure polyene chain of $N+2$ members. We have then, for HLPS

$$\Delta E=4\alpha\sin\frac{3}{2}\frac{\pi}{N+3}\sin\frac{\pi}{2(N+3)}$$

$$-4\frac{e-f}{N+3}\sin^2\frac{\pi}{N+3}\left\{4\cos^2\frac{\pi}{N+3}\right.$$

$$-1+\frac{1}{R}\frac{N}{N+3}\cos\frac{\pi}{N+3}$$

$$\left.\times\left(4\cos\frac{\pi}{N+3}\cos\frac{2\pi}{N+3}-1\right)\right\}. \tag{36}$$

Now the wave-lengths of the absorption band go to infinity with $(N+3)^2$ with a correction proportional to $N+3$ (just as for $R=1$).

For LCAO

$$\Delta E=2\alpha\left(\sin\frac{\pi}{N+3}\right)$$

$$\times\left\{1+\frac{N}{(N+3)^2}\frac{2}{R^2}\cos^2\frac{\pi}{N+3}\right\}$$

$$-\frac{4(e-f)}{N+3}\sin^2\frac{\pi}{N+3}. \tag{37}$$

There we have again, for large N, the same behavior as for $R\leqq1$.

4. For values of R and N, for which our approximations are not good, numerical calculations have been performed. E, and therefore k or v, were assumed and the corresponding R calculated from (17), (18), (17′). The results were plotted and selected values read, which are given in Table I.

5. Change in interaction energy of the ends. We can also treat the case in which the ends of the ion have a different interaction energy, α', with the C atoms than the C atoms among themselves. Then (8) remains unchanged, while (6) and (7) are replaced by

$$(E-e)X=+\alpha' x_1, \tag{6′}$$

$$(E-f)x_1=\alpha' X+\alpha x_2 \tag{7′}$$

with corresponding changes in (9) and (10). For the lowest states, (14′) and (15′) remain unchanged, while the energy equation (replacing

(17′)) takes the form

$$1-\frac{v}{R}=\gamma v^2\pm v^{N+1}\frac{1-v^2}{(1\pm\gamma)^{N+1}} \qquad (17''')$$

with

$$\gamma=1-\frac{\alpha^2}{\alpha'^2}.$$

Similarly (17) and (18) are replaced by

$$2\cos k-\frac{1}{R}=\frac{\alpha'^2}{\alpha^2}\frac{\cos\frac{N-1}{2}k}{\cos\frac{N+1}{2}k} \qquad (18'')$$

and a corresponding equation with sinus, (13), (14), (15) remaining unchanged. The calculation shall be limited to the two lowest states, given by (17″) for $R<1$. It is better to divide (17‴) by v^2. One then finds

$$\frac{1}{v}=\frac{1}{2R}[1+(1+4\gamma R^2)^{\frac{1}{2}}]\pm v^{N-1}(1-v^2)\frac{R}{(1+4\gamma R^2)^{\frac{1}{2}}}.$$

Call[12]

$$\frac{(1+4\gamma R^2)^{\frac{1}{2}}-1}{2\gamma R}=\epsilon.$$

Then

$$\frac{1}{v}=\frac{1}{\epsilon}\pm\frac{R}{(1+4\gamma R^2)^{\frac{1}{2}}}\epsilon^{N-1}(1-\epsilon^2)$$

and

$$v=\epsilon\pm\frac{R}{(1+4\gamma R^2)^{\frac{1}{2}}}\epsilon^{N+1}(1-\epsilon^2).$$

Therefore

$$E=f+\alpha\left(\frac{1}{\epsilon}+\epsilon\right)\pm\frac{R\alpha}{(1+4\gamma R^2)^{\frac{1}{2}}}\epsilon^{N-1}(1-\epsilon^2)^2 \qquad (20')$$

and the energy difference of the first transition according to HLPS

$$\Delta E=\frac{2R\alpha}{(1+4\gamma R^2)^{\frac{1}{2}}}(1-\epsilon^2)^2\epsilon^{N-1}. \qquad (21')$$

(21′) has the same form as (21), only the constants are different, ϵ replacing R and $2R\alpha/\epsilon(1+4\gamma R^2)^{\frac{1}{2}}$ replacing 2α. The coefficients of the different structures in each state, according to HLPS, are given by (13), (14), (13′), and (16). In the LCAO interpretation these are the coefficients of the atomic orbitals in the molecular orbital. In both cases, these coefficients are not yet normalized.

[12] It is assumed that $|\gamma R|<1$, but γ may be negative.

TABLE I. $\Delta E/\alpha$, accordingly to HLSP.

R	$N=5$	$N=7$	$N=9$
0.70	0.105		0.021
0.75	0.120	0.058	0.032
0.80	0.135	0.072	0.044
0.85	0.17	0.085	0.052_5
0.90	0.18	0.100	0.061
0.95	0.19	0.105	0.074
1.00	0.20	0.120	0.081
1.5	0.28	0.185	0.135
2.0	0.325	0.220	0.160
2.5	0.345	0.235	0.170
10	0.410	0.270	0.195

The transition moments have been calculated as $\sim N+2$ by Mulliken[13] (for polyenes).

III. THE ENDS ARE UNEQUAL

In this case we again assume the same interaction energy between X and x_1, as between Y and x_n, but take the energies of X and Y to be e_1 and e_2 instead of e. (10) has to be changed accordingly. The states are now not anymore purely symmetric or antisymmetric.

Therefore we try to integrate (8) by

$$x_s=av^{[(N+1)/2]-s}+bv^{-[(N+1)/2]+s}, \qquad (38)$$

which takes the place of (13) and (14) if $v=e^{ik}$ and the place of (13′) if v is real. The difference between a and b expresses the fact that the two progressive waves which make up the standing wave (13) and (14) are not any more equal here.

Insertion of (38) into (8) leads again to (15′). Then, (6) and (7) give

$$E-f=\frac{\alpha^2}{E-e_1}+\alpha\frac{x_2}{x_1}=\frac{\alpha^2}{E-e_1}+\alpha\frac{av^{[(N+1)/2]-2}+bv^{-[(N+1)/2]+2}}{av^{[(N+1)/2]-1}+bv^{-[(N+1)/2]+1}}. \qquad (39)$$

Similarly, (9) and (10) result in

$$E-f=\frac{\alpha^2}{E-e_2}+\alpha\frac{x_{N-1}}{x_N}=\frac{\alpha^2}{E-e_2}+\alpha\frac{av^{[(N+1)/2]-(N-1)}+bv^{-[(N+1)/2]+(N-1)}}{av^{[(N+1)/2]-N}+bv^{-[(N+1)/2]+N}}. \qquad (40)$$

[13] R. S. Mulliken, J. Chem. Phys. **7**, 20, 364 (1939).

Eliminate b/a between these two equations, and the following equation is found from a rather lengthy calculation:

$$v^{2N-2}\left\{\frac{E-f}{\alpha}+\frac{\alpha}{E-e_1}-\frac{1}{v}\right\}\left\{\frac{E-f}{\alpha}\frac{\alpha}{E-e_2}-\frac{1}{v}\right\}$$
$$=\left\{\frac{E-f}{\alpha}+\frac{\alpha}{E-e_1}-v\right\}\left\{\frac{E-f}{\alpha}+\frac{\alpha}{E-e_2}-v\right\}.$$

Using again (15′) one finds

$$v^{2N+2}(1-vR)(1-vR')=\left\{1-\frac{R}{v}\right\}\left\{1-\frac{R'}{v}\right\}, \quad (41)$$

where

$$R=\frac{\alpha}{e_1-f}, \quad R'=\frac{\alpha}{e_2-f}.$$

For $R=R'$, this reduces to (19), as one can easily show. We call the lower structure $e_1(e_1>e_2$, $R<R')$.

1. If R, $R'\ll 1$, we find, as in II, two solutions with real v. In this case[14] zero approximation gives (by setting the left side zero)

$$v=R, \qquad v=R'. \quad (42)$$

This inserted in the left side of (41), leads to

$$v=\tfrac{1}{2}(R+R')-\tfrac{1}{2}\{(R'-R)^2$$
$$+4R^{2N+4}(1-R^2)(1-RR')\}^{\frac{1}{2}}, \quad (43)$$

$$v'=\tfrac{1}{2}(R+R')+\tfrac{1}{2}\{(R'-R)^2$$
$$+4R'^{2N+4}(1-R'^2)(1-RR')\}^{\frac{1}{2}}. \quad (43')$$

One can with the same approximation, write $1/v$ and $1/v'$. Abbreviate

$$\sigma^2=4R^{2N}(1-R^2)(1-RR') \quad (44)$$

and call σ' the result of exchanging R' and R in σ. One finds then for the lowest levels

$$E_1=\frac{e_1+e_2}{2}+\frac{\alpha}{2}(R+R')+\tfrac{1}{2}\{(e_1-e_2)^2+\alpha^2\sigma^2\}^{\frac{1}{2}}$$
$$-\tfrac{1}{2}RR'\left\{(e_1-e_2)^2+\alpha\sigma^2\frac{R^2}{R'^2}\right\}^{\frac{1}{2}}, \quad (45)$$

$$E_2=\frac{e_1+e_2}{2}+\frac{\alpha}{2}(R+R')-\tfrac{1}{2}\{(e_1-e_2)^2$$
$$+\alpha^2\sigma'^2\}^{\frac{1}{2}}+\tfrac{1}{2}RR'\left\{(e_1-e_2)^2+\alpha^2\sigma'^2\frac{R'^2}{R^2}\right\}^{\frac{1}{2}}. \quad (45')$$

Therefore, the split is

$$\Delta E=\tfrac{1}{2}\{(e_1-e_2)^2+\alpha^2\sigma^2\}^{\frac{1}{2}}$$
$$+\tfrac{1}{2}\{(e_1-e_2)^2+\alpha^2\sigma'^2\}^{\frac{1}{2}}$$
$$-\tfrac{1}{2}RR'\left\{(e_1-e_2)^2+\alpha^2\sigma^2\frac{R^2}{R'^2}\right\}^{\frac{1}{2}}$$
$$-\tfrac{1}{2}RR'\left\{(e_1-e_2)^2+\alpha^2\sigma'^2\frac{R'^2}{R^2}\right\}^{\frac{1}{2}}. \quad (46)$$

In (43), (43′), (45), (45′), (46) the absolute value of the square root is meant. If N is sufficiently big so that $(e_1-e_2)^2>\alpha^2\sigma^2(R^2/R'^2)\ldots$ (which is more stringent than the condition for the applicability of (43), if e_1-e_2 is small compared to $e-f$), the formula can be simplified to

$$\Delta E=(e_1-e_2)(1-RR')$$
$$+\frac{\alpha^2}{4}\frac{1}{e_1-e_2}\left\{\sigma^2\left(1-\frac{R^2}{R'}\right)+\sigma'^2\left(1-\frac{R'^2}{R}\right)\right\}, \quad (46')$$

which for $N=\infty$, approaches the finite value $(e_1-e_2)(1-RR')$. If R, R' are larger than 1, and for the higher levels of the case where R, R' are smaller than 1, one has to substitute for v in (38) e^{ik} and gets then

$$RR'\sin(N+3)k-(R+R')\sin(N+2)k$$
$$+\sin(N+1)k=0. \quad (47)$$

$k=0$ and $k=\pi$, which formally satisfy (47), are not solutions of the problem.[15] This also is true for the corresponding value $v=1$ in Eq. (38).

(47) takes the place of the two Eqs. (17), (18). We now solve again separately for the two cases:

$$R<1, \quad R'<1 \qquad \text{and} \qquad R>1, \quad R'>1.$$

[14] The numerical limits within which the approximation is good are the same as given in footnote 8.

[15] From these values of k (or $v=1$) follows that x_s would be independent of s. Then Eqs. (6), (7) give

$$(E-f)=\frac{\alpha^2}{E-e_1}+\alpha$$

while (9), (10) lead to

$$E-f=\frac{\alpha^2}{E-e_2}+\alpha,$$

which cannot both be true for $e_1\neq e_2$.

In the former, we would have, for $R=R'=0$, $\sin(N+1)k=0$, or $k=r\pi/(N+1)$. We therefore set

$$k=\frac{r\pi}{N+1}+\delta, \quad r=1\cdots N \tag{48}$$

and

$$RR'\sin\left\{\frac{2r\pi}{N+1}+\frac{N+3}{N+1}\delta\right\}$$

$$-(R+R')\sin\left\{\frac{r\pi}{N+1}+\frac{N+2}{N+1}\delta\right\}+\sin\delta=0. \tag{49}$$

If N is very large one can write, to the first approximation

$$\operatorname{tg}\delta=\left[R+R'-2RR'\cos\frac{r\pi}{N+1}\right]$$

$$\times\left(\sin\frac{r\pi}{N+1}\right)\left\{1-(R+R')\cos\frac{r\pi}{N+1}\right.$$

$$\left.+RR'\cos\frac{2r\pi}{N+1}\right\}^{-1}. \tag{50}$$

The highest filled state, in the LCAO method, is again

$$r=\frac{N-1}{2}$$

or

$$\operatorname{tg}\delta_1=\left\{R+R'-2RR'\sin\frac{\pi}{N+1}\right\}$$

$$\times\left(\cos\frac{\pi}{N+1}\right)\left\{1-(R+R')\sin\frac{\pi}{N+1}\right.$$

$$\left.-RR'\cos\frac{2\pi}{N+1}\right\}^{-1}$$

or

$$\operatorname{tg}\delta_1\sim\left\{R+R'-2RR'\frac{\pi}{N+1}\right\}$$

$$\times\left\{1-RR'-(R+R')\frac{\pi}{N+1}\right\}^{-1} \tag{51}$$

or even more roughly

$$\operatorname{tg}\delta_1=\frac{R+R'}{1-RR'}. \tag{51'}$$

For the first empty state $r=(N+1)/2$

$$\operatorname{tg}\delta_2=(R+R')(1-RR')^{-1}, \tag{51''}$$

which, to this approximation, agrees with $\operatorname{tg}\delta_1$.

The energy of the transition is then

$$\Delta E=2\alpha\cos\left[\frac{N-1}{N+1}\frac{\pi}{2}+\delta_1\right]$$

$$-2\alpha\cos\left[\frac{N+1}{N+1}\frac{\pi}{2}+\delta_1\right]$$

$$=2\frac{\alpha\pi}{N+1}(1-RR')(1+R^2)^{-\frac{1}{2}}(1+R'^2)^{-\frac{1}{2}}. \tag{52}$$

This formula is correct for sufficiently large N, without demanding that R and R' have to be very small compared to 1.

Therefore, for LCAO we find here, just as for equal ends, that the wave-length of the first transition goes to infinity as $N+1$.

2. Instead of the special case of $R=1$ of II, we have here the special case $RR'=1$ which can be solved exactly. Then (47) takes the form

$$(\sin(N+2)k)\left[2\cos k-\left(R+\frac{1}{R}\right)\right]=0 \tag{53}$$

with the solutions:

$$k=r\pi/(N+2) \qquad r=1\cdots N+1 \tag{53'}$$

and, in addition, the lowest lying term, given by $v=R$ (Eq. 42); $v=R'=R^{-1}$ describes the same state. Therefore the two lowest states are

$$E_1=f+\alpha\left(R+\frac{1}{R}\right)=f+(e_1+e_2-2f) \tag{54}$$

and

$$E_2=f+2\alpha\cos\frac{\pi}{N+2}. \tag{54'}$$

The first transition in the HLSP approximation is

$$\Delta E=\alpha\left[R+\frac{1}{R}-2\cos\frac{\pi}{N+2}\right] \tag{55}$$

and approaches, for $N=\infty$, the finite value

$$\Delta E_\infty=\alpha\left[R+\frac{1}{R}-2\right]=e_1+e_2-2f-2\alpha$$

$$=[(e_1-f)^{\frac{1}{2}}-(e_2-f)^{\frac{1}{2}}]^2. \tag{55'}$$

For a LCAO description, the highest filled state is given by $r=(N-1)/2$, the lowest empty one by $r=(N+1)/2$. These have the energy

$$E_1=f+2\alpha\cos\frac{N-1}{N+1}\frac{\pi}{2}=f+2\alpha\sin\frac{\pi}{N+1}, \quad (56)$$

$$E_2=f+2\alpha\cos\frac{N+1}{N+1}\frac{\pi}{2}=f \quad (56')$$

and the energy of the first transition is

$$E=2\alpha\sin\frac{\pi}{N+1} \quad (57)$$

with a wave-length going to infinity proportional to $N+1$, as for $R=1$ in a symmetrical ion.

3. For $R>1$, $R'>1$ we rewrite (47) as

$$\sin(N+3)k-\left(\frac{1}{R}+\frac{1}{R'}\right)\sin(N+2)k+\frac{1}{RR'}\sin(N+1)k=0 \quad (47')$$

and solve by putting

$$k=\frac{r\pi}{N+3}+\delta, \quad r=1\cdots N+2 \quad (58)$$

or

$$\sin\delta-\left[\frac{1}{R}+\frac{1}{R'}\right]\sin\left\{\frac{N+2}{N+3}\delta-\frac{r\pi}{N+3}\right\}+\frac{1}{RR'}\cdot\sin\left\{\frac{N+1}{N+3}\delta-\frac{2r\pi}{N+3}\right\}=0. \quad (59)$$

For large N, one finds now

$$\sin\delta=-\left[\left(\frac{1}{R}+\frac{1}{R'}\right)-\frac{2}{RR'}\cdot\cos\frac{r\pi}{N+3}\right]\times\left(\sin\frac{r\pi}{N+3}\right)\left[1-\left(\frac{1}{R}+\frac{1}{R'}\right)\cos\frac{r\pi}{N+3}+\frac{1}{RR'}\cos\frac{2r\pi}{N+3}\right]^{-1}. \quad (59')$$

Therefore, for ground and excited state in HLPS ($r=1, 2$),

$$\delta_1=-\left\{\frac{1}{R}+\frac{1}{R'}-\frac{2}{RR'}\right\}\times\left\{1-\frac{1}{R}-\frac{1}{R'}+\frac{1}{RR'}\right\}^{-1}\cdot\frac{\pi}{N+3}, \quad (60)$$

$$\delta_2=-2\delta_1, \quad (60')$$

$$\Delta E=2\alpha\cos\left[\frac{\pi}{N+3}+\delta_1\right]-2\alpha\cos\left[\frac{2\pi}{N+3}+2\delta_1\right]=3\alpha\left(\frac{\pi}{N+3}\right)^2\left\{1-\left(\frac{1}{R}+\frac{1}{R'}-\frac{2}{RR'}\right)\times\left[1-\frac{1}{R}-\frac{1}{R'}+\frac{1}{RR'}\right]^{-1}\right\}^2. \quad (61)$$

Therefore the wave-length goes to infinity like $(N+3)^2$ as for equal ends.

If the LCAO gives the better description, then we are interested in the transition between the terms with $r=(N+1)/2$ and $r=(N+3)/2$.

The corresponding values of δ are, in the zero approximation, for both levels

$$\operatorname{tg}\delta=-\frac{\frac{1}{R}+\frac{1}{R'}}{1-\frac{1}{RR'}}. \quad (62)$$

The levels themselves are

$$E_1=f+2\alpha\sin\left[\frac{\pi}{(N+3)}-\delta\right], \quad (63)$$

$$E_2=f-2\alpha\sin\delta \quad (63')$$

and the transition energy

$$\Delta E=\frac{2\alpha\pi}{N+3}\cos\delta=\frac{2\alpha\pi}{N+3}\left(1-\frac{1}{RR'}\right)\times\left(1+\frac{1}{R^2}\right)^{-\frac{1}{2}}\left(1+\frac{1}{R'^2}\right)^{-\frac{1}{2}}. \quad (64)$$

This is identical with the expression (52), which gives ΔE according to LCAO for $R<1$,

$R'<1$, except that the absolute value is to be taken and that we have $N+3$ instead of $N+1$, which difference has no significance for the present accuracy.

4. Finally there exists the possibility

$$R<1 \qquad R'>1 \qquad RR'\neq 1.$$

Then the lowest state is given by (41), with v near R, or in the first approximation

$$v=R+R^{2N+1}\frac{RR'-1}{R'-R}(1-R^2),$$

$$\frac{1}{v}=\frac{1}{R}-R^{2N+2}\frac{RR'-1}{R'-R}(1-R^2),$$

$$E_1=f+\alpha\left\{R+\frac{1}{R}-R^{2N+2}\frac{RR'-1}{R'-R}(1-R^2)^2\right\}. \quad (65)$$

The higher states are contained in (47), which, divided by $R'+R$, can be written

$$\sin(N+2)k-R\frac{R'}{R'+R}\sin(N+3)k-\frac{1}{R'+R}\sin(N+1)k=0. \quad (66)$$

The approximate solutions are

$$(N+2)k=r\pi+\delta \qquad r=1\cdots N+1, \quad (67)$$

of which there are the proper number, $N+1$.

Therefore

$$\sin\delta=R\frac{R'}{R'+R}\sin\left\{\frac{r}{N+2}\pi+\frac{N+3}{N+2}\delta\right\}-\frac{1}{R'+R}\sin\left\{\frac{r\pi}{N+2}-\frac{N+1}{N+2}\delta\right\}. \quad (68)$$

For large N or small R, $1/R'$

$$\sin\delta=\frac{RR'-1}{R'+R}\sin\frac{r}{N+2}\pi, \quad (68')$$

$$E=f+2\alpha\left\{\cos\frac{r\pi}{N+2}-\frac{1}{N+2}\frac{RR'-1}{R'+R}\sin^2\frac{r\pi}{N+2}\right\}. \quad (69)$$

Therefore, according to HLPS, the first absorption corresponds to the energy

$$\Delta E=E_1-E_2=-2\alpha\cos\frac{\pi}{N+2}+\alpha\left(R+\frac{1}{R'}\right)+\alpha\left\{-R^{2N+2}\frac{(1-R^2)^2}{R'-R}+\frac{1}{N+2}\cdot\left(\frac{1}{R'+R}\right)\sin^2\frac{\pi}{N+2}\right\}(RR'-1), \quad (70)$$

while the LCAO method gives (transition from $r=(N-1)/2$ to $r=(N+1)/2$)

$$\Delta E=4\alpha\cos\frac{\pi}{2(N+2)}\sin\frac{\pi}{N+2}+\frac{1}{N+2}\frac{RR'-1}{R+R'}\left[\sin^2\frac{3\pi}{2(N+2)}-\sin^2\frac{\pi}{2(N+2)}\right]. \quad (71)$$

Therefore ΔE approaches, for increasing N, a finite limiting value, if one or both of the R's are smaller than 1 and if the HLPS method is used, in all other cases it goes to zero.

IV. THE NEUTRAL MOLECULE[16]

The neutral base of the dye has two ground structures, one of energy e_1.

$$H_2N-(CH=CH)_n-CH=NH \quad (72)$$

and one of energy e_2

$$H_2\overset{+}{N}=CH-(CH=CH)_n-\overset{-}{N}H, \quad (73)$$

and N excited structures of energy f

$$NH_2-\overset{+}{C}H-(CH=CH)_n-\overset{-}{N}H. \quad (74)$$

Formally, therefore, the situation is exactly the same as the one treated in III. For R, $R'\ll 1$, the transition energy is given by (46), (46') in HLPS and by (52) in the HHM method.

For $RR'=1$, it is (55) for HLPS and (57) for LCAO.

For R, $R'>1$, it is expressed by (61) in the HLPS method and by (64) according to LCAO.

[16] See also U. Fano, Phys. Rev. **61**, 736(A) (1942).

Finally, if $R<1$, $R'>1$ and $RR'\neq 1$, ΔE is found from (70) for HLPS and from (71) for LCAO.

Just as for the unsymmetrical ion, ΔE goes to a finite limit with increasing N only if at least one of the R is smaller than 1 and the HLPS method is used, in all other cases it goes to zero.

V. A DIFFERENT ASSUMPTION ABOUT THE ION [17]

In the preceding paragraphs, we have given to the structures of type (2) and type (3) the same energy f. It is possible that it would be a better numerical approximation to take the energy of (I) and (3) equal, namely e. In this case, the structures have alternately energies e and f, $(N+3)/2$ having e, $(N+1)/2$ the energy f. The equivalent differential equations are

$$\ddot{x}_1=-ex_1-\beta x_2, \tag{75}$$

$$\ddot{x}_2=-fx_2-\beta(x_3+x_1), \tag{76}$$

$$\ddot{x}_{2s}=-fx_{2s}-\beta(x_{2s-1}+x_{2s+1}), \tag{77}$$

$$\ddot{x}_{2s+1}=-ex_{2s+1}-\beta(x_{2s}+x_{2s+2}), \tag{78}$$

$$\ddot{x}_{N+2}=-ex_{N+2}-\beta x_{n+1}. \tag{79}$$

For integration, one writes

$$x_{2s}=a\{v^{[(N+3)/2]-2s}\pm v^{-([(N+3)/2]+2s)}\}, \tag{80}$$

$$x_{2s+1}=b\{u^{[(N+3)/2]-2s-1}\pm u^{-([(N+3)/2]+2s+1)}\}. \tag{81}$$

Writing the Eqs. (77), (78) in the form

$$E-f=\beta\frac{x_{2s-1}+x_{2s+1}}{x_{2s}}, \tag{77'}$$

$$E-e=\beta\frac{x_{2s}+x_{2s+2}}{x_{2s+1}}, \tag{78'}$$

the insertion of (80), (81) and the requirement that the right side be independent of s, leads to

$$u=v.$$

Then (77'), (78') take the form

$$E-f=\beta\frac{b}{a}\left(v+\frac{1}{v}\right), \tag{82}$$

$$E-e=\beta\frac{a}{b}\left(v+\frac{1}{v}\right) \tag{82'}$$

or

$$(E-f)(E-e)=\beta^2\left(v+\frac{1}{v}\right)^2, \tag{83}$$

which replaces (15), (15').

Furthermore, writing (75) and (76) in the analogous forms to (77'), (78') and multiplying leads to

$$(E-f)(E-e)=\beta^2\frac{x_3+x_1}{x_1}. \tag{84}$$

The use of (80) and (83) then gives

$$\frac{v^{[(N+3)/2]+1}\pm v^{-([(N+3)/2]+1)}}{v^{[(N+3)/2]-1}\pm v^{-([(N+3)/2]-1)}}=-1. \tag{85}$$

This has no solution for a real $v<1$. Putting therefore again $v=\exp.(ik)$ one finds, for the symmetric solutions,

$$\cos\left(\frac{N+5}{2}k\right)=-\cos\frac{N+1}{2}k \tag{86}$$

and for the antisymmetric ones

$$\sin\frac{N+5}{2}k=-\sin\frac{N+1}{2}k, \tag{87}$$

which leads to

$$k=\frac{r\pi}{N+3}, \tag{88}$$

r odd for symmetric, even for antisymmetric solutions.

A short transformation shows that

$$x_j=\left.\begin{matrix}a\\b\end{matrix}\right\}\sin\left(j\frac{r\pi}{N+3}\right), \tag{80'}$$

where the coefficient is a or b, depending on the question whether j is odd or even.

One sees that $r=1\cdots N+2$, which at first glance seems to give the correct number of states.

The energy follows from (83) as

$$E_r=\frac{e+f}{2}\pm\frac{1}{2}\left\{(e-f)^2+16\beta^2\cos^2\frac{r\pi}{N+3}\right\}^{\frac{1}{2}}. \tag{89}$$

We have now to investigate the sign in (89). Physically it seems clear that the positive sign should go with $(N+3)/2$ states, which lie near

[17] By mistake, the letter β has been used in this chapter for the quantity previously called α.

e for small β, the negative sign with the $(N+1)/2$ states lying near f.

But from the mathematical standpoint, the double sign seems to give, at first glance, two energy values for each state.

We see clearer, if we remember that x_j, i.e., the coefficients in the wave function, is not defined by the *sine* alone, but also by the ratio a/b.

Compare the state r with the state $N+3-r=r'$.

We find

$$x_j'=\left.\begin{matrix}a\\b\end{matrix}\right\}\sin\left[j\frac{\pi}{N+3}(N+3-r)\right]$$

$$=\begin{matrix}a\\b\end{matrix}(-1)^{j-1}\sin\frac{\pi rj}{N+3}. \quad (90)$$

That is to say: If a/b has the same sign in r and r', the phase of every alternate "particle" differs, in the state r', by 180° from the phase of the same "particle" in the state r (analogous to the "optical branch" in crystal vibrations). If, however, a/b has the opposite sign for r and r', the alternate "particles" have the same phase in r and r' (acoustical branch). In other words, what matters is the ratio

$$a\sin\frac{\pi rj}{N+3}:b\sin\frac{\pi r}{N+3}(j+1),\quad j=2s+1.$$

In the energy formula (89), $\cos 2\pi r/(N+3)$ has the *same* value for r and r'.

It follows from (75), (89) that

$$\beta\frac{x_2}{x_1}=E-e=\frac{f-e}{2}$$

$$\pm\frac{1}{2}\left\{(e-f)^2+16\beta^2\cos^2\frac{r\pi}{N+3}\right\}^{\frac{1}{2}}. \quad (91)$$

Therefore, with the positive sign, x_2/x_1 is positive; with the negative sign, negative. To sum up:

r need only be taken from 1 to $(N+3)/2$.

For $r=(N+3)/2$, we have only one solution, namely (see 82′)

$$E=e \qquad a=0 \qquad \text{(positive sign in (89)).}$$

The negative sign in (89), giving $E=f$, to which would correspond $b=0$, does not satisfy Eq. (82). To $E=e$ belongs the wave function

$$x_{2s+1}=(-1)^s$$
$$x_{2s}=0.$$

For each r which obeys

$$1\leqq r\leqq\frac{N+1}{2},$$

there are two solutions.

The positive sign in (89) gives an energy larger than e. The corresponding wave functions have the character of the "acoustical branch."

The negative sign in (89) gives an energy smaller than f, the corresponding wave functions resembling the "optical branch."

The energy difference between the two lowest energy values is

$$\Delta E=\frac{1}{2}\left[\left\{(e-f)^2+16\beta^2\cos^2\frac{\pi}{N+3}\right\}^{\frac{1}{2}}-\left\{(e-f)^2+16\beta^2\cos^2\frac{2\pi}{N+3}\right\}^{\frac{1}{2}}\right]$$

$$=\frac{1}{2}\left[\left\{(e-f)^2+16\beta^2-16\beta^2\sin^2\frac{\pi}{N+3}\right\}^{\frac{1}{2}}-\left\{(e-f)^2+16\beta^2-16\sin^2\frac{2\pi}{N+3}\right\}^{\frac{1}{2}}\right], \quad (92)$$

which, for large N, takes the approximate form

$$\Delta E=4\beta^2\{(e-f)^2+16\beta^2\}^{-\frac{1}{2}}\frac{3\pi^2}{(N+3)^2}. \quad (92')$$

This investigation was started at the suggestion of A. L. Sklar and has profited by many discussions with him.

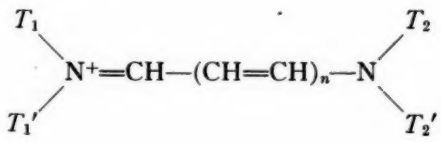
Energy Levels and Color of Polymethine Dyes

A. L. SKLAR
Chemical Laboratory, The Catholic University of America, Washington, D. C.

(Received April 29, 1942)

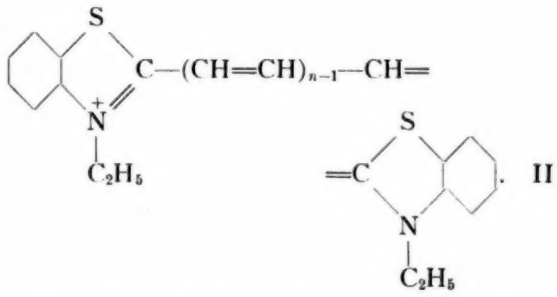
The secular determinant for the energy levels of the unsaturation electrons of a polymethine dye is discussed in both HLSP and LCAO approximations. The roots of the secular determinant, which were obtained in the preceding paper by Herzfeld, are applied to a discussion of the longest wave-length electronic band of symmetrical and unsymmetrical polymethines. The LCAO approximation gives good numerical values for the dependence of both the transition energy and oscillator strength on the length of the polymethine chain in symmetrical dyes, but cannot handle the questions which depend sensitively on a small difference between the groups attached to the two nitrogen atoms at the extreme ends of the chain. Although the HLSP method yields correct qualitative results for symmetrical dyes, it does not give good numerical values. The HLSP method, however, is sensitive to a difference between the two ends of the dye molecule and affords a simple explanation for a number of properties in which unsymmetrical dyes differ from symmetrical ones.

THE spectra of the polymethine dyes have been discussed briefly from the atomic orbital point of view by Pauling[1] and from the molecular orbital viewpoint (LCAO) by Mulliken[2] and Förster.[3] The connection between the molecular structure and the low lying energy levels of these dyes will be discussed qualitatively in a forthcoming paper;[4] here it will be treated more quantitatively in order to arrive at a firmer basis for the purely qualitative considerations. The lower energy levels of the plane ion I will be approximated by both the atomic orbital (HLSP) and molecular orbital (LCAO) methods. Ion I represents a large class of polymethine dyes, if the T groups are



I

understood to mean either individual groups or a nucleus which bends around and joins to the conjugated chain, as in the ion II.



II

Since the excitation of the in plane σ electrons yields the molecular Rydberg series below 2000A, we may restrict our discussion concerning the visible and near ultraviolet absorption to the energy values of the unsaturation electrons (π or P_z) which do not interact with the σ electrons.

In the LCAO method, as is well known, we write a linear combination of atomic orbitals for the molecular orbital of each unsaturation electron in the field of the multiply-charged residue formed by stripping ion I of its unsaturation electrons. Ion I has $(2n+4)$ unsaturation electrons but we will consider that two of these are permanently localized, one on each nitrogen atom. Our system, then, consists of $(2n+2)$ electrons in a potential field made up of a chain of $(2n+3)$ singly charged nuclei. If we write the one-electron orbitals as a linear combination of the $(2n+3)$ atomic orbitals and neglect all overlaps and also interactions between electrons on non-neighboring atoms, the energies of a single electron are

[1] L. Pauling-Gilman, *Organic Chemistry*, Vol. 2, p. 888.
[2] R. S. Mulliken, J. Chem. Phys. **7**, 570 (1939).
[3] T. Förster, Zeits. f. physik. Chemie **B47**, 245 (1940); **48**, 12 (1941).
[4] A. L. Sklar and L. G. S. Brooker (to be published).

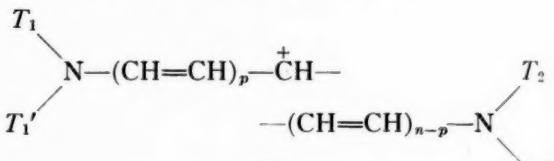
given by the roots of the $(2n+3)$ rowed secular determinant III, whose solution has been discussed in the preceding paper.[5]

$$\begin{vmatrix} e_1-W; & \alpha_1 ; & 0 ; & 0; & \cdots 0; & 0 ; & 0 ; & 0 \\ \alpha_1 ; & f-W; & \alpha ; & 0; & \cdots 0; & 0 ; & 0 ; & 0 \\ 0 ; & \alpha ; & f-W; & \alpha; & \cdots 0; & 0 ; & 0 ; & 0 \\ 0 ; & 0 ; & \alpha ; & \diagdown & \vdots & & & \\ \vdots & \vdots & \vdots & & & & & \\ 0 ; & 0 ; & 0 ; & \cdots & \alpha; & f-W; & \alpha ; & 0 \\ 0 ; & 0 ; & 0 ; & \cdots & 0; & \alpha ; & f-W; & \alpha_2 \\ 0 ; & 0 ; & 0 ; & \cdots & 0; & 0 ; & \alpha_2 ; & e_2-W \end{vmatrix} = 0. \qquad \text{III}$$

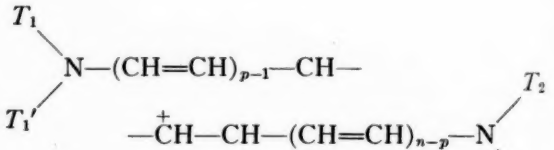
In this determinant W is the one-electron energy; α is the interaction energy between two electrons on neighboring carbon atoms; α_1 and α_2 are, respectively, the interactions between two electrons on a carbon atom and on either the left- or right-hand nitrogen atom; and f, e_1, and e_2 are, respectively, the ionization potentials of an unsaturation electron localized on either a carbon atom, the left- or right-hand nitrogen atom. In this approximation the f's will be somewhat different for the various carbon atoms in the chain. It will, however, be sufficiently accurate for our purposes to take all the f's as equal,[6] especially since the variation in f in this approximation is due primarily to the fact that the interaction between the unsaturation electrons is neglected. In the ground state of the ion I the $(2n+2)$ electrons will fill the lowest $(n+1)$ orbitals and the longest wave-length electronic band will be associated with the transition of an electron from the $(n+1)$st to the $(n+2)$nd orbital.

In going over to the Heitler-London-Slater-Pauling method it is to be noticed that the ground structure I is doubly degenerate in that the positive charge may be shifted to the other nitrogen atom and the double bonds readjusted. Ion I is thus analogous to benzene which has the two Kekulé structures as doubly degenerate ground structures. A difficulty, however, arises in carrying over from benzene to ion I the view that the longest wave-length electronic absorption is associated with a transition between the two molecular states which arise from a resonance splitting of the degenerate ground structures.[7] The electron distributions in the two I structures, which differ in that an electron has moved from one end of the molecule to the other, do not overlap appreciably. The split of these two structures would thus be expected to be negligible instead of the observed value of the order of a few volts.

Pauling,[8] however, has pointed out that one must consider, in addition to the two structures I, the following set of intermediate structures $X_{2p+1}(p=0 \text{ to } n)$. Since the overlap of electrons on atoms separated by even two interatomic distances is very small, we will also have to include the set of structures $Y_{2p}(p=1 \text{ to } n)$



X_{2p+1}



Y_{2p}.

[5] K. F. Herzfeld, J. Chem. Phys. **10**, 508 (1942).

[6] We also neglect energy differences due to the replacement by other groups of the hydrogen atoms attached to the conjugated carbon chain.

[7] A. L. Sklar, J. Chem. Phys. **5**, 669 (1937); H. Sponer, G. Nordheim, A. L. Sklar, and E. Teller, J. Chem. Phys. **7**, 207 (1939).

[8] L. Pauling, Proc. Nat. Acad. Sci. **25**, 577 (1939).

For every structure in the set consisting[9] of I, X, and Y there is a second structure which differs from the former in that an electron has been moved to the neighboring atom on the left. Thus, although the two I structures do not overlap appreciably directly, one I structure interacts with X_1, X_1 with Y_2, Y_2 with X_3, and so on until X_{2n+1} interacts with the other I structure. The two ground structures I can split through this chain of interactions with the intermediate structures X and Y.

Although it is not easy to estimate numerically the relative energy of I and X, it is clear from chemical considerations[10] that the I structures are considerably lower in energy than the X and Y structures. The Y structures are energetically higher than the X structures since the former have one less double band. Acting against this effect is the additional possibility of resonance possessed by the Y structures due to the fact that the C^+ breaks the carbon chain into two odd segments which may be written several ways, as for example:

$$-(CH{=}CH)_k{-}CH{-}\rightleftarrows{-}CH{-}(CH{=}CH)_k{-}.$$

To each Y, then, corresponds a group of structures and we may consider that, in writing one structure for this group, we have written the stabilized result of the resonance among the whole group. Since two compensating factors enter in the relative energy of the X and Y structures, and since a small difference in their energy will not affect our discussions, we will assume, for mathematical simplicity, that all intermediate structures, X and Y, are of equal energy.

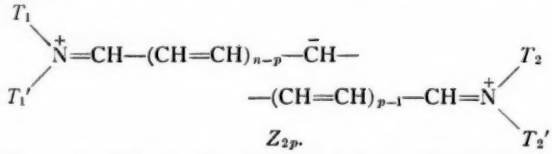
In all we have $2n+3$ structures, two I, $(n+1)$X, and nY structures. If we neglect all Coulomb and exchange integrals which are smaller than the product of the first power of the overlap integral of electrons on neighboring atoms times the ordinary Coulomb and exchange integrals, respectively, and also neglect terms containing overlaps of electrons on non-neighboring atoms, the molecular energies resulting from resonance among the $2n+3$ structures are given by the roots of a secular determinant which is formally the same as that reached in the LCAO method, namely determinant III. The meaning of the quantities in the determinant is, of course, now quite different. α, α_1, α_2 are, respectively, the interaction energies between an intermediate structure, and either a second intermediate or one of the two ground structures; e_1 and e_2 are the energies of the two ground structures I; and f is the energy of an intermediate structure, X and Y. Since W is the energy of all the unsaturation electrons in the molecule, the transition energy in question is the difference between the two smallest values of W.

The roots of determinant III have been discussed as a function of the number of rows in the preceding paper by Herzfeld[5] for various ranges of the parameters $R=\alpha/(e_1-f)$ and $R'=\alpha/(e_2-f)$. The results will now be applied to a discussion of the qualitative conclusions which were reached in reference 4, since purely qualitative considerations of complex situations, even when applied with caution, are occasionally misleading.

TABLE I.

	HLSP		LCAO	
$R\ll 1$	$2\alpha(1-R^2)R^N$	(21)	$2\alpha\sin\frac{\pi}{N+1}$	(26)
$R=1$	$4\alpha\sin^2[\pi/2(N+2)]$	(30)	$4\alpha\cos\frac{\pi}{N+2}\sin\frac{\pi}{2(N+2)}$	(31)
$R\gg 1$	$4\alpha\sin\frac{3}{2}\frac{\pi}{N+3}$	(36)	$2\alpha\sin\frac{\pi}{N+3}$	(37)

[9] The set of structures X and Y are chosen in preference to the following set, $Z_k(k=1$ to $2n+1)$, on energetic grounds.



Z_{2p}.

It is not obvious, though, that the interaction energy among the Z structures is not sufficiently larger than that among the X, Y structures to compensate for the higher energy of the former. This question is, however, immaterial for the purpose of this paper since we will not try to evaluate the interaction energy but only its variation as n is increased or as the character of the T groups is changed.

[10] The essence of the chemical arguments is the fact that whereas ammonium ions N^+ are frequently encountered, carbonium ions C^+ are never met in any quantity, although they probably do occur in minute amounts in the course of reactions. As an example, we may compare the basicity of ammonia ($NH_3+H_2O\rightarrow NH_4^++OH^-$) with that of ethylene ($CH_2{=}CH_2+H_2O\rightarrow CH_3{-}CH_2^++OH^-$). Except for the very small difference between the strength of a CH and NH bond, the second reaction requires more energy than the first by about the same amount that is required to change structure I into an X structure. (The hydration energies balance out to a first approximation.) The fact that ethylene is enormously weaker as a base than ammonia would, then, suggest that X is considerably higher in energy than I, since there is no reason to suspect the entropy to be vastly different in the two cases.

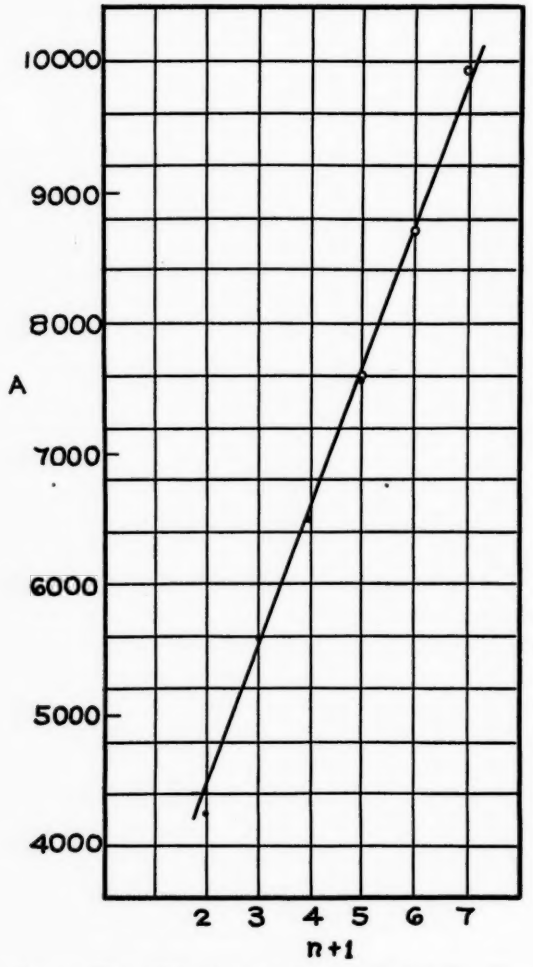
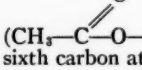


FIG. 1. The solid points in the figure give the wavelength of the absorption peak for ion II as a function of the number of double bonds between the nitrogen atoms (NH). The empty circles give the same data for an ion which differs from ion II only in having an acetoxy group



in place of the hydrogen atom on the sixth carbon atom of the polymethine chain.

SYMMETRICAL IONS

We will first treat those dyes in which the same groups are attached to both nitrogen atoms as, for example, in ion II. For such symmetrical ions the secular determinant is simplified since $\alpha_1=\alpha_2$ and $e_1=e_2$. Since α and α_1 are of the same order of magnitude and since a small difference between them would not be expected to affect the general questions treated in this paper, except for very small values of n, we will further simplify the mathematical treatment by setting α_1 equal to α.

Table I gives the transition energies to the first order found in reference 5 from the roots of the secular determinant, for various ranges of the parameter $R=\alpha/(e-f)$. In Table I the numbers in parentheses are the numbers of the equations in reference 5 and $N=(2n+1)$ is the number of carbon atoms in the chain between the two bounding nitrogen atoms.

It is clear from the formulas in Table I that the magnitude of R plays a determining role in the HLSP but not in LCAO approximation, except for very small values of N. This becomes clear, physically, if we regard α as constant and vary R by varying $e-f$. In the HLSP method the energy of the degenerate ground structures, e is the zero approximation to the energies of the two lowest states of the molecules. Therefore, the transition energy depends on the magnitude of the split of the degenerate energies of the ground structures. Since these two structures interact only through a chain of second-hand interactions with the intermediate structures X and Y, which are higher in energy than the former by an amount $(f-e)$, it is clear that[11] the greater is $(f-e)$, the smaller is the transition energy. Table I of reference 5, obtained by a numerical solution of Eqs. 17 and 18, shows this dependence if α is considered to be constant; the transition energy increases as $(f-e)$ decreases (R increasing) and approaches for small values of $(f-e)$ the limiting value, independent of R, required by Eq. 36.

In the LCAO method, on the other hand, e is the zero approximation, not to the energy of the molecule, but to the lowest energy state of one electron of the molecule. Since the states associated with e, as well as a number of states whose energies in zero approximation are f, are full, both in the ground and excited states, there is no longer any reason to expect $(f-e)$ to play a large role in the transition energy.

It is clear from Table I that in both approximations and for all ranges of R, the transition energy should decrease monotonously with increasing chain length N as is found to be the case.[12]

[11] $f-e$ may be changed experimentally by changing the character of the groups T_j which are attached to the nitrogen atom. For example, the ionization energy of a electron from methyl amine, $(CH_3)\cdot NH_2$ is smaller than that of aniline, $(C_6H_5)_2NH$.

[12] L. G. S. Brooker *et al.*, J. Am. Chem. Soc. **62**, 1116 (1940).

The algebraic relationship between the peak wave-length and N differs in the two methods and, in the HLSP approximation, depends on R. The usual linear dependence of peak wave-length on chain length shown in Fig. 1[13] agrees with the results found for large values of N by the LCAO approximation for all ranges of R. In the HLSP approximation the dependence on N is quadratic for $R \gtrsim 1$, but is exponential for very small values of R; only in the region of R slightly smaller than unity is there a possibility of a linear dependence.

The value of the absorption frequency for fixed N is directly proportional to the interaction energy α for constant R, but the ratio of the frequencies of a dye with a given value of N and that with an N larger by one, is not explicitly dependent on α. The value of this ratio affords a test for the two methods. In the LCAO method, the ratio does not depend sensitively on the value of R and is therefore completely determined by N; in HLSP approximation, the ratio does depend on R. We can compare the theoretical values with the experimental[14] value of 1.17 (cf. Fig. 1) for the ratio of the peak absorption of a dye with a chain of seven carbon atoms ($n=3$) to that of one with a chain of nine carbon atoms ($n=4$). The LCAO approximation, in excellent agreement with experiment, yields the value 1.19 for this ratio when R is very large; 1.20 when R is unity, and 1.24 for very small values of R. In the HLSP approximation, however, the above ratio varies between 1.2 and 1.4 as R varies between one-half and ten.[15]

In addition to yielding too high a value for the above ratio, the HLSP method demands an abnormally high value for the exchange energy α to explain the absolute values. The transition energy for a five-membered carbon chain ($n=2$), which is actually about three volts, is only two-tenths α in the region $R=1$, which seems to be the most likely region in this method. In the LCAO approximation the transition energy, roughly independent of R, for a five-membered carbon chain, is about nine-tenths α, which is much more reasonable.

It is seen that although both methods tell us that the absorption frequency should decrease as the chain lengthens, the quantitative features of this decrease are not given very well by the HLSP method, but are given very well indeed by the LCAO approximation.

We may also inquire into the dependence of the intensity on chain length. Although the chain is a bent one, we may assume that each increase of n in ion I by unity lengthens the chain by a constant amount. The coefficients of the orbitals in the LCAO approximation are given, except for a normalization factor, in Eqs. 13 and 14 of reference 5. With them we can calculate the transition moment, which, as Mulliken[2] has pointed out, is polarized along the chain. In first order, the transition moment turns out to be proportional to N when N is large. Since the intensity is proportional to the product of the frequency and the square of the transition moment, and since the frequency in the LCAO approximation is inversely proportional to N, the intensity should vary linearly with N. The intensity of the shorter members of the series represented by ion II (small values of n) is shown in Table II.[16] Although the first member is out of line, the last three $\mathbf{f}_0$ values vary linearly with the chain length as indicated by the last two successive differences. Since the theoretical discussion on the variation of intensity with chain length was limited to large values of N, the excellent agreement of the last three values is probably in part fortuitous. It would, accordingly, be useful to obtain $\mathbf{f}$ values for the higher members of this series but these, unfortunately, are not very stable and rapidly change with time.

In addition to yielding the correct dependence

TABLE II.

N	$\mathbf{f}_0$
3	1.2_2
5	1.2_2
7	1.5_6
9	1.8_9

$\mathbf{f}_0$ is the oscillator of the longest wave-length electronic band of ion II and $N=2n+1$ is the number of carbon atoms in the chain.

[13] The data in Fig. 1 were obtained at the Eastman Kodak Laboratories and communicated to the author by Dr. L. G. S. Brooker.

[14] It has been pointed out (reference 12) that the wave-length of the peak of most symmetrical polymethine ions increases by about 1000 angstrom units when n increases by one. The constancy of this value is to be expected according to Table I except in the HLSP approximation when $R \ll 1$.

[15] Reference 5, Table I.

[16] The data in Table II were calculated from the absorption curves of Fig. 2.

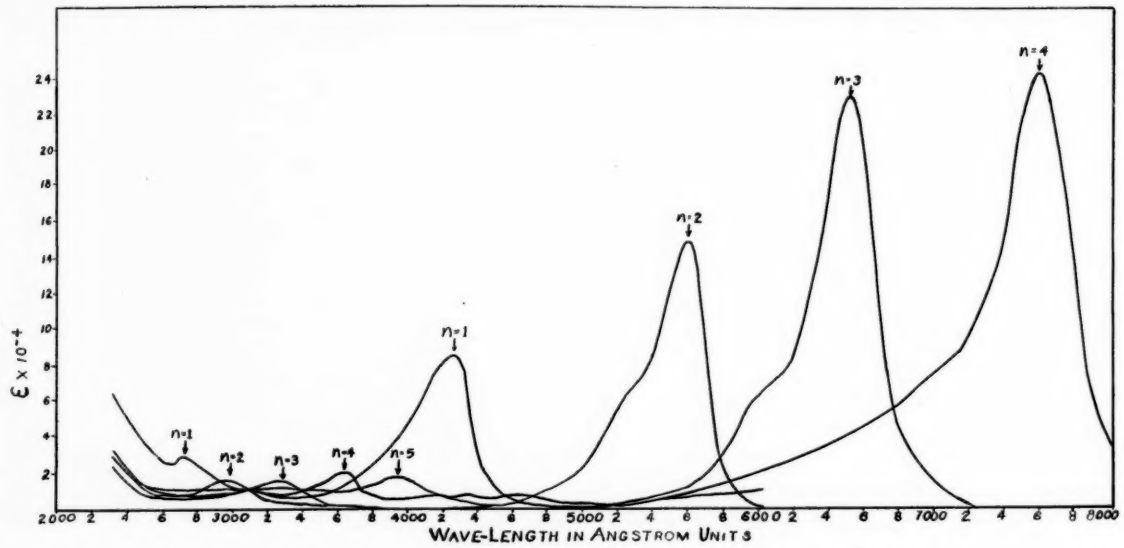


FIG. 2. The absorption curves of ion II for $n=1$ to 4; $N=3$ to 9.

of intensity on chain length, the LCAO method also has yielded interesting results concerning the relative intensities of the set of absorption bands of a given dye. On the basis of the LCAO method, Mulliken has reported[2] that in a series of $N \rightarrow V$ transitions, as are our transitions, in a maximally elongated chain, the longest wave-length transition is by far the strongest. Since the ions II have bulky nuclei on the ends, their chains may be expected to be in an elongated configuration. Figure 2 shows,[17] in accord with Mulliken's prediction, that the higher energy transitions are very weak compared with the lowest energy transition. For example, the oscillator strength of the band having a peak at 3600A ($N=9$) is 0.2_3, the long wave band having an oscillator strength of 1.8_9.

UNSYMMETRICAL IONS

If the nuclei T_2 which are attached to the right-hand nitrogen atom in ion I, differ from those attached to the left-hand one, the ionization energies of the two nitrogen atoms are no longer equal and e_1 is, in general, different from e_2 in both methods. In the HLSP approximation, we would expect a sensitive dependence of the color on the difference between the two e's since the resonance splitting of the energies of the two I structures is both affected by, and superimposed on, the energetic difference (e_1-e_2). In the LCAO approximation, on the other hand, this difference clearly plays a minor role in the transition energy. The special sensitiveness of the HLSP method to (e_1-e_2) is due to the fact that in this method e_1, e_2 are the zero-order approximations to the energies of the two levels involved in the transition; whereas, in the LCAO approximation, the two one-electron levels approximated in zero order by e_1, e_2 are filled in both the ground and excited state of the molecule.

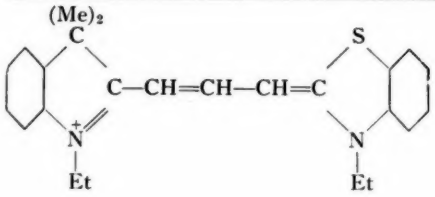
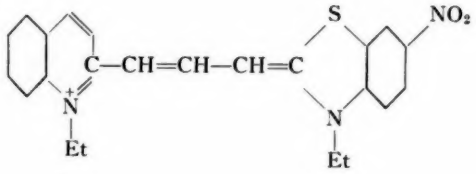
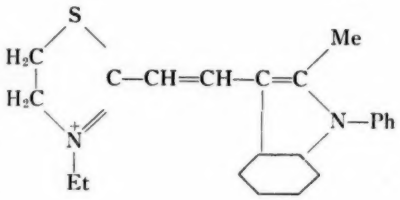
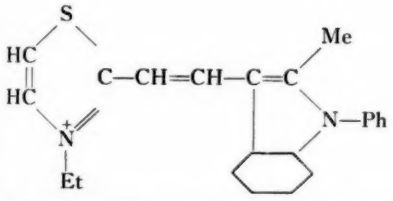
That differences in kind between symmetrical and unsymmetrical ions do exist may be seen experimentally by comparing the absorption peak of an unsymmetrical ion with the mean of the absorption peaks of the two parent symmetrical ions of which the unsymmetrical ion may be considered a hybrid. The existence of important differences between the absorption peak of unsymmetrical ions and the mean of those of its two parents has been emphasized by Brooker[18] and called the "deviation." Isolated examples[18] of the deviations are given in Table III.

As discussed above we can only hope to get an explanation of the significant differences between unsymmetrical and symmetrical ions through the HLSP approximation. However, before any confidence can be placed in results which are yielded by the HLSP but not by the LCAO method, one

[17] The absorption curves of Fig. 2 were taken at the Eastman Kodak Laboratory and communicated to the author by Dr. L. G. S. Brooker.

[18] L. G. S. Brooker and Sprague, J. Am. Chem. Soc. 63, 3202 (1941).

TABLE III. $\tilde{\nu}$ is the wave number of the absorption peak of the ion in column I, $\tilde{\nu}_L(\tilde{\nu}_R)$ is that for the symmetrical ion having the nucleus on the left (right) in common with the unsymmetrical ion, and the last column is the deviation. All ions occurring in this table have $N=5$. Me is a CH_3 group, Et a C_2H_5 group, and Ph a phenyl group. It is to be noted that in the ions given in the last two lines of the table the end groups containing conjugated double bonds are joined to an inner member of the chain. It is not clear how far the present formulae apply to this case. However, the table shows that the rules discussed do apply and that the deviation is particularly large.

Ion	$\tilde{\nu}$	$\tilde{\nu}_L$	$\tilde{\nu}_R$	Dev.
	18,150	17,710	17,930	+330
	17,240	16,560	17,140	+390
	23,950	22,470	22,270	+1580
	22,730	17,510	22,270	+2840

must ask whether the results are artefacts of the HLSP approximation, especially since the LCAO method gave better numerical results on symmetrical dyes.

That one can trust, in a qualitative manner, the conclusions which the HLSP method gives concerning the effect of the difference (e_1-e_2) can be seen by comparing the direction of the errors made by the two methods. As discussed before, e_1, e_2 affect the transition energy sensitively in HLSP because they are zero-order approximations to the energies of the two molecular states involved in the transition; whereas, in the LCAO approximation, the difference (e_1-e_2) has no appreciable effect on the molecular states. The true molecular state, however, is somewhere between those described by the two methods, being somewhat closer to the LCAO as judged by the numerical results concerning the dependence of the transition energy on chain-length. Since the actual situation is intermediate between that described by the two methods, the approximation which overemphasizes a given effect is certainly more reliable for finding, in a qualitative way, the consequences of that effect than is the approximation which practically neglects the effect in question.[19]

Our discussion of unsymmetrical ions, $e_1 \neq e_2$,

[19] Since, in the HLSP approximation, e_1, e_2 are important primarily because the two structures I have more chemical bonds, one might hope that the dependence of the molecular energies on (e_1-e_2) will reappear in the molecular orbitals method if antisymmetrical molecular orbitals, including spin, are used and the interaction of electrons explicitly included in the Hamiltonian. See M. Goeppert Mayer and A. L. Sklar, J. Chem. Phys. **6**, 645 (1938).

will accordingly be limited to the HLSP approximation and the results will depend on the magnitude of R.

The same factors which cause e_1 to differ from e_2 also introduce a difference between the interaction energies α_1 and α_2. However, a distinction between the roles played by an inequality of e_1 and e_2 and that played by an inequality of α_1 and α_2 appears if one compares an unsymmetrical dye with the two parent symmetrical dyes of which the unsymmetrical dye may be considered a hybrid. Both ground structures are degenerate of energy e_1 in one parent symmetrical dye and of energy e_2 in the other parent; whereas, in the unsymmetrical dye the ground structures are no longer degenerate, but have the energies e_1 and e_2 in first order. Since the e's are zero approximations to the energies of the molecular states involved in the transition, we might anticipate, as will be shown later, that the transition energy of the unsymmetrical dye will be larger than that of either parent symmetrical dye. That is, on the basis of a difference in e_1 and e_2, we should expect the unsymmetrical dye to differ from any mean of its parent symmetrical dyes. On the other hand, if e_1 were equal to e_2, even if α_1 were not equal to α_2, we would expect the properties of the unsymmetrical dye to be intermediate between those of its symmetrical parents. Further theoretical investigation will be required to decide which type of mean one should expect if α_1 differed from α_2 but if e_1 were essentially equal to e_2.

The properties, then, which are peculiar to unsymmetrical dyes, and which distinguish them in kind from symmetrical dyes, may be found experimentally by comparing the absorption peak[20] of the unsymmetrical dye with those of the symmetrical parents, and may be found theoretically by setting $\alpha_1=\alpha_2$ but keeping e_1 different from e_2.

We will, then, set α_1 equal to α_2,[21] but keep e_1 different from e_2, in the secular determinant III, and thereby obtain a transition energy for the unsymmetrical dye which is to be contrasted with a mean of the transition energies of its two symmetrical parents. The results depend upon the order of magnitude of $R=\alpha/(e_1-f)$ and $R'=\alpha/(e_2-f)$.

Case I. $R\ll 1$ and $R'\ll 1$

The easiest case to discuss qualitatively is that of small values of R and R' which is the case when the intermediate structures, X and Y, are considerably higher in energy than the ground structures I; the qualitative discussions of reference 4 were limited to this case. Equation 46 of reference 5 gives the transition energies when R and R' are small as:

$$2\Delta E_N=\{(e_1-e_2)^2+(\alpha\sigma_N)^2\}^{\frac{1}{2}}+\{(e_1-e_2)^2+(\alpha\sigma_N')\}^{\frac{1}{2}}, \quad (46)$$

where $\sigma_N^2=4R^{2N}(1-R^2)(1-RR')$.

σ_N' has R and R' interchanged. $\alpha\sigma_N$, $\alpha\sigma_N'$ differ from the transition energies of the parent symmetrical dyes only in that the factor $(1-RR^1)$ replaces a factor $(1-R^2)$ which is a difference of higher order for the case of small R under discussion.

A number of interesting results follow from Eq. 46. First, it is clear that the effect of a small non-degeneracy (e_1-e_2) causes the transition energy ΔE_N to be larger than $(\sigma_N+\sigma_N')\alpha/2$, which is essentially the mean of the transition energies of the parent symmetrical dyes. Thus, it is seen that unsymmetrical dyes should absorb at shorter wave-lengths, that is, be "lighter" in color, than related symmetrical dyes. This has been found by Brooker[22] who, in a series of papers, gives values of the "deviation," which are always found to be toward short waves except in one or two cases when they are extremely small. Examples of the deviation are shown in Table III.

Although the transition energy is larger, the resonance stabilization in the ground state is for unsymmetrical dyes smaller than the mean of their related parent dyes. This may easily be seen from Eqs. 20 and 45 of reference 5, by comparing the difference between the energy of the lower ground structure and that of the ground state for the various cases. One should accordingly expect that unsymmetrical dyes, by virtue of their larger resonance stabilization, should be

[20] Since the absorption curves are obtained in solution the significance of the wave-length of peak absorption is open to question. This can perhaps be settled by a study, now under way of the structure of the absorption bands at low temperatures.

[21] As in the case of symmetrical dyes, we also equate α_1 and α.

[22] L. G. S. Brooker *et al.*, J. Am. Chem. Soc. **62**, 1116 (1940); *ibid.*, **63**, 3192 (1941); *ibid.*, **63**, 3203 (1941); *ibid.*, **63**, 3214 (1941); *ibid.*, **64**, 199 (1942).

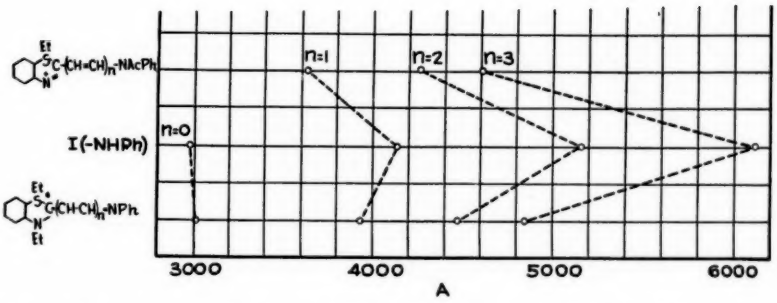
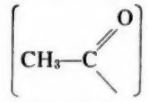


FIG. 3. "Converging" and "non-converging" series. The dye in the center of the figure differs from the one at the top only in that the acetate radicle



is replaced by a hydrogen atom.

more stable than unsymmetrical dyes. They should, for example, be less readily decomposed.

An interesting confirmation has been called to the author's attention by Brooker. The cyanine dyes act as indicators both in an acid and alkaline range. Now, either the addition or subtraction of a proton will be resisted by the resonance stabilization of the cyanine chain since the resonance is present only before either of these processes. On this basis of a larger resonance stabilization of symmetrical dyes, one should expect the applicability to the polymethine dyes of a rule which has been found by Schwarzenbach[23] and which states that the difference between the pK values[24] (pH at half-transformation) is larger for symmetrical than for unsymmetrical dyes.

It may also be seen from Eq. 46 that as N increases, σ_N, and accordingly ΔE_N, decreases. Now in a symmetrical dye ΔE_N would go to zero as N is increased indefinitely, but in an unsymmetrical dye, ΔE_N decreases to the magnitude of the energy difference (e_1-e_2). It has, indeed, been found by Brooker[22] that for many series of unsymmetrical dyes the members of which differ only in the value of N, the absorption peaks appear to converge to a finite wave-length limit as N increases, whereas the peaks for a corresponding series of symmetrical dyes do not converge to a finite line when plotted on a wave-length scale. Many unsymmetrical dyes form series which behave like series of symmetrical dyes; the former presumably have small values of the difference (e_1-e_2). Figure 3 shows an example of a "non-converging" series (middle case) and two examples of "converging" series. The dye at the bottom of Fig. 3 is not an ion but its analogy to the present case is discussed in the last section.

Equation 46 also affords an explanation of the "sensitivity rule" found by Brooker,[22] which states that the deviation produced by introducing a given alteration in the molecular structure of a symmetrical molecule is much smaller than the change in deviation produced by making the same structural alteration in a highly unsymmetrical molecule. This may be seen to be a consequence of the fact that the resonance interaction and the energetic difference (e_1-e_2) enter into the transition energy as a sum of squares. When the difference (e_1-e_2) is small compared to $\alpha\sigma_N$, a given increase in (e_1-e_2) affects ΔE_N much less than would the same increase if (e_1-e_2) were of the order of magnitude of $\alpha\sigma_N$ or larger. One set of examples from Brooker's papers is given in Table IV to illustrate the "sensitivity rule."

Case II. $RR'=1$

In order to see whether the results discussed in Case I are valid even if the intermediate structures are not very much higher in energy than the ground structures, we may consider an unsymmetrical dye in which $RR'=1$. This case can be solved exactly; Eq. 55 of reference 5 gives the following transition energy when $RR'=1$:

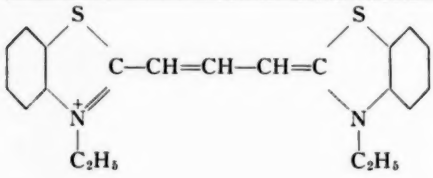
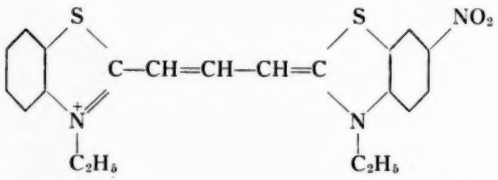
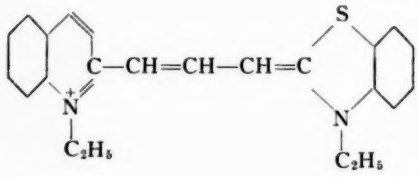
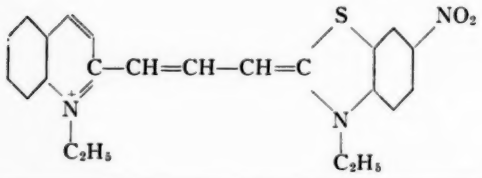
$$\Delta E_N=\alpha\left\{R+R'-2\cos\frac{\pi}{N+2}\right\}. \qquad (55)$$

One should expect, even in Case II, a finite convergence limit for the absorption peaks of a

[23] G. Schwarzenbach, Zeits. f. Electrochemie **47**, 40 (1941). Schwarzenbach's examples are not strictly analogous to our case but will be shown in a later paper to have the same basic explanation.

[24] Since the difference in the pK's is a measure of the free energy change involved in the reaction $\{2I\rightarrow HI^{++}+B\}$ where B is the neutral molecule obtained by taking a proton away from ion I, one would not expect the entropy change in the case of unsymmetrical dyes to differ markedly from that of symmetrical dyes, except perhaps for an effect due to the symmetry number. The symmetry number comes into question because both HI^{++} and B may be unsymmetrical whether I is symmetrical or unsymmetrical. This effect, however, is at most small (of the order of a few tenths of a pH unit), and even in the wrong direction.

TABLE IV. $\tilde{\nu}$ is the wave number of the absorption peak in methyl alcohol of the dye in column I, $\tilde{\nu}_L$, $\tilde{\nu}_R$ are those of the parent symmetrical dye, which have both ends like the left (right) side of the ion in column I; Dev. is the deviation and ΔD is the change in deviation produced by introducing a nitro group. All ions in Table IV have five CH groups between the nitrogen atoms. Note that the ion in the second row differs from that in the first row only in that a hydrogen atom is replaced by a nitro group; the fourth ion differs from the third in exactly the same way, but because the third ion is unsymmetrical (deviation 55), the effect of the introduction of the NO_2 is much greater than for the symmetrical ion in the first line.

Ion	$\tilde{\nu}$	$\tilde{\nu}_L$	$\tilde{\nu}_R$	Dev.	ΔD
	17,930	—	—	—	
	17,500	17,930	17,140	35	35
	17,300	16,560	17,930	55	
	17,240	16,560	17,140	390	345

homologous series of dyes as N increases indefinitely, though a smaller one than in Case I. This may be seen from equation 55′ which applies to Case II.

$$\lim_{N\to\infty} \Delta E_N = \{(e_1-f)^{\frac{1}{2}}-(e_2-f)^{\frac{1}{2}}\}^2. \quad (55')$$

For values of (e_1-e_2) which are small compared to (e_1-f) this limit reduces to:

$$(e_1-e_2)^2/4(e_1-f),$$

which is smaller than (e_1-e_2), (Case I).

The finite convergence limit implies that one should expect that unsymmetrical ions, even in Case II, should exhibit a "deviation," at least for large values of N.

That the "sensitivity rule" should hold even in Case II may also be seen by considering the convergence limit rather than the deviation. The energy difference (e_1-e_2) is a function of e_1 since $RR'=1$. If we now differentiate the convergence limit $(R+R'-2)$ with respect to e_1, we find that the derivative is proportional to the energy difference (e_1-e_2). That is, we should expect that there exists a "sensitivity rule" for convergence limit which would seem to imply the same for the deviation, at least for large values of N.

The discussions concerning resonance stabilization cannot be treated in the same way as in the preceding section because only the symmetrical case in which $R=1$ has been solved exactly. Since the cases when R is slightly above or below unity have not been treated exactly, it is not simple to compare the behavior of the unsymmetrical ion, $RR'=1$, with the mean behavior of the two symmetrical parent ions.

That the resonance stabilization is less for unsymmetrical than for symmetrical ions, even in Case II, can be seen, however, in the following

way, at least when R does not differ too much from R'. In this case, although the behavior of an unsymmetrical ion cannot be compared with the mean behavior of its symmetrical parents, it can be compared to the behavior of a symmetrical ion having an R equal to the mean of those of its symmetrical parents since if $RR'=1$ and the R's are almost equal, the mean of R and R' is unity. The resonance stabilization, E_1-e_1, is equal to α in the symmetrical case, $R=1$ (cf. Eq. 54, reference 5), and equal to $R'\alpha$ in the unsymmetrical case, $RR'=1$ (cf. Eq. 54, reference 5) where R' is the smaller of the two R's and hence less than unity. Since the resonance stabilization of unsymmetrical ions is less than that of symmetrical ions when R and R' are not very different, one would expect this same to be true, *a fortiori*, when the R's do differ considerably.

Case III. $R\gg1$; $R'\gg1$

In case both R and R' are large, the intermediate structures are close to the ground structures and one should expect this case to differ in kind from Case I. From chemical considerations[10] one would not expect this case to obtain in the polymethine dyes.

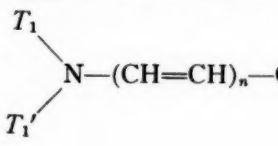
Equation 61 of reference 5 gives the transition energy in Case III as:

$$\Delta E_N=3\alpha\left(\frac{\pi}{N+3}\right)^2\left\{1-\left(\frac{1}{R}+\frac{1}{R'}-\frac{2}{RR'}\right)\times\left(1-\frac{1}{R}-\frac{1}{R'}+\frac{1}{RR'}\right)^{-1}\right\}^2. \quad (61)$$

In this case, Eq. 61 shows that there is no finite convergence limit to the wave-length as N increases indefinitely. This is not surprising since the ground structures, whose energetic difference caused the convergence limit in Cases I and II, are now very bad approximations to the two lowest states of the molecule.

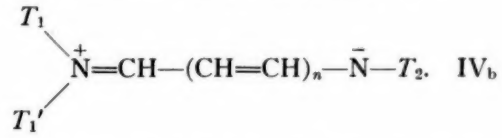
Case IV. $R\ll1$, $R'\gg1$

This case is useful in considering the neutral dye molecule in which the lowest structure is IV_a:



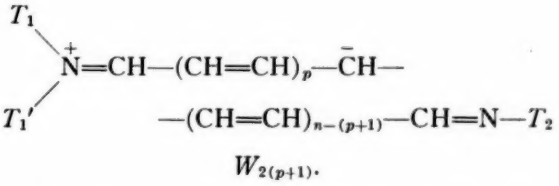
$$T_1,\ T_1'\text{—}N\text{—}(CH{=}CH)_n\text{—}CH{=}N\text{—}T_2. \quad IV_a$$

A number of intermediate structures exist of which the lowest in energy is probably IV_b:



$$T_1,\ T_1'\text{—}\overset{+}{N}{=}CH\text{—}(CH{=}CH)_n\text{—}\bar{N}\text{—}T_2. \quad IV_b$$

The remaining $(2n+1)$ intermediate structures are:



$$T_1,\ T_1'\text{—}\overset{+}{N}{=}CH\text{—}(CH{=}CH)_p\text{—}\bar{C}H\text{—}(CH{=}CH)_{n-(p+1)}\text{—}CH{=}N\text{—}T_2$$

$$W_{2(p+1)}.$$

Since the set of $(2n+3)$ structures IV_a, IV_b, and W_j interact among themselves in the same way as the set of $(2n+3)$ structures for ion I, the neutral dye molecule may be considered as a limiting case of a highly unsymmetrical ion in which one of the ground structures, IV_a, is of very much lower energy than the intermediate structures, but the other ground structure IV_b is just a little lower in energy than the intermediate structures, $W_{2(p+1)}$. This case is handled, then, by treating R as small and R' as large compared to unity. The results have already been discussed for the special case that $RR'=1$. Equation 70 of reference 5 gives the transition energy for Case IV as:

$$\frac{1}{\alpha}\Delta E_N=\left(R+\frac{1}{R'}\right)-2\cos\frac{\pi}{N+1}+\left\{\frac{(R+R')}{N+2}\sin^2\frac{\pi}{N+2}-R^{2N+2}\frac{(1-R^2)^2}{R'-R}\right\}(RR'-1). \quad (70)$$

As N increases indefinitely, ΔE_N for the neutral molecule approaches a finite limit just as in the case of a highly unsymmetrical ion.

In conclusion the author would like to express his appreciation for numerous discussions with Professor K. F. Herzfeld and Dr. L. G. S. Brooker, and also for the courtesies shown him by the Eastman Kodak Laboratories.

A Practical Method for the Solution of Certain Problems in Quantum Mechanics by Successive Removal of Terms from the Hamiltonian by Contact Transformations of the Dynamical Variables

Part I. General Theory

L. H. THOMAS
Mendenhall Laboratory of Physics, The Ohio State University, Columbus, Ohio

(Received December 18, 1941)

Problems in quantum mechanics may be solved by canonical change of the representation in terms of which the dynamical variables are expressed. They may also be solved by contact transformations of the dynamical variables. The usual notation is modified slightly to make it more convenient for developing these two methods. Their parallel development shows their essential equivalence but formal difference.

INTRODUCTION

THE motion of a molecule is complicated and requires for its description a number of dynamical variables which may be chosen in various ways.[1] If we choose variables, such as the normal coordinates for small oscillations, which vary with the time nearly independently, a zero-order approximation to the motion is obtained by neglecting entirely the interaction between these variables.[2] If the electric moment and polarizability of the molecule are expressed in terms of them, the absorbtion, emission, and scattering of light may be calculated.[3] The gross structure of the spectrum of the molecule can be explained in this approximation, but to explain the details, the interactions must be taken into account. This is conveniently done by perturbation theory or successive approximation, going to the first order, second order, or even to higher orders, in the small interactions.[4] This more accurate work is necessary also if we wish to obtain the forces between the atomic nuclei in the molecule from the observed spectrum with even moderate precision.[5]

In the customary method perturbation theory is applied to the representation of the motion in terms of the stationary states of the zero-order approximation.[6] When there are many variables this becomes cumbersome, and is liable to error because of the large number of separate terms to be considered. We have found in a number of problems that the work is considerably abbreviated by applying transformations directly to the normal coordinates; and this liability to error is reduced.[7] While this is one of the earliest methods of quantum mechanics and is explained in many textbooks, it was found less useful than the customary method in dealing with atomic spectra and it has been little used.[8] The transformations are parallel to the contact transformations used in dynamical astronomy to remove terms from the Hamiltonian function.

1. NOTATION

The usual notations are a little clumsy for dealing with contact transformations, and we modify them slightly as follows:

A symbol in square brackets as

$$[s] \tag{1.01}$$

[1] H. Kronig, *Band Spectra and Molecular Structure* (Cambridge, 1930); H. Sponer, *Molekulspectren II* (Springer, 1936); G. Herzberg, *Molecular Spectra and Molecular Structure I*, (Prentice Hall, 1939); D. M. Dennison, Rev. Mod. Phys. **3**, 280 (1931); **12**, 175 (1940).

[2] H. Kronig, reference 1, p. 31; H. Sponer, reference 1, p. 42.

[3] H. Kronig, reference 1, Chapter 3; G. Herzberg, reference 1, Chapter 3.

[4] Born, Heisenberg, and Jordan, Zeits. f. Physik **35**, 557 (1925).

[5] D. M. Dennison, reference 1, p. 177.

[6] A. Ruark and H. C. Urey, *Atoms, Molecules and Quanta* (McGraw-Hill, 1930), p. 597; P. A. M. Dirac, *Quantum Mechanics*, second edition, (Oxford, 1934), Chapter 8; E. C. Kemble, *Fundamental Principles of Quantum Mechanics* (McGraw-Hill, 1937), Chapter 11.

[7] Shaffer, Nielsen, and Thomas, Phys. Rev. **56**, 895 (1939); W. H. Shaffer, J. Chem. Phys. **9**, 607 (1941); S. Silver and W. H. Shaffer, J. Chem. Phys. **9**, 599 (1941).

[8] A. Ruark and H. C. Urey, reference 6, p. 595; P. A. M. Dirac, reference 6, p. 110; E. C. Kemble, reference 6, p. 373.

will stand for a matrix with elements

$$(m|s|u). \tag{1.02}$$

In this symbol m and u are variables or sets of variables designating the row and column of the matrix; they may be given by several variables taking integral values, or continuous ranges of values, or both. In general there need be no correlation implied between the values of m and the values of u, but most usually u is given by corresponding variables to those giving m taking the same values, and we then conveniently denote them by m and n or by n and n'.

The physical systems in which we are interested are described in terms of dynamical variables, ξ, η, $\cdots$, which combine according to a non-commutative algebra we suppose known, given by kinematical conditions which usually comprise commutation relations: e.g., a coordinate q and its canonically conjugate momentum p satisfying

$$qp-pq=ih \tag{1.10}$$

(h, as usual will stand for Planck's constant, 6.62×10^{-27} erg-sec.; $\hbar$ stands for $h/2\pi$).

The dynamical variables can be represented accordingly in various ways by Hermitian matrices,

$$[a|\xi|a],\ [a|\eta|a],\ \cdots, \tag{1.11}$$

with elements

$$(a_m|\xi|a_n),\ (a_m|\eta|a_n),\ \cdots. \tag{1.12}$$

In these symbols m and n are the variables or sets of variables designating the row and column of the matrix; a designates the representation; and ξ, η, $\cdots$, designate the dynamical variable which the matrix represents. Since $[a|\xi|a]$ is Hermitian, $(a_m|\xi|a_n)$ and $(a_n|\xi|a_m)$ are conjugate complex numbers.

E.g., if $q\sqrt{(2\mu\omega/\hbar)}$ and $p\sqrt{(2/\mu\omega\hbar)}$ are represented by the matrices

$$\begin{array}{cccccc} 0 & i & 0 & 0 & 0 & \cdot\cdot \\ -i & 0 & i\sqrt{2} & 0 & 0 & \cdot\cdot \\ 0 & -i\sqrt{2} & 0 & i\sqrt{3} & 0 & \cdot\cdot \\ 0 & 0 & -i\sqrt{3} & 0 & i\sqrt{4} & \cdot\cdot \\ 0 & 0 & 0 & -i\sqrt{4} & 0 & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \end{array}$$

and

$$\begin{array}{cccccc} 0 & 1 & 0 & 0 & 0 & \cdot\cdot \\ 1 & 0 & \sqrt{2} & 0 & 0 & \cdot\cdot \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \cdot\cdot \\ 0 & 0 & \sqrt{3} & 0 & \sqrt{4} & \cdot\cdot \\ 0 & 0 & 0 & \sqrt{4} & 0 & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \end{array}$$

and if we call the representation E, then $[E|q\sqrt{(2\mu\omega/\hbar)}|E]$ and $[E|p\sqrt{(2/\mu\omega\hbar)}|E]$ are these matrices, and if we number the rows from the top by $m=0, 1, 2, \cdots$, and the columns from the left by $n=0, 1, 2, \cdots$,

$$\begin{aligned} (E_m|q\sqrt{(2\mu\omega/\hbar)}|E_n) &= i\sqrt{(\tfrac{1}{2}m)} && n=m+1 \\ & && m, n=0, 1, 2, \cdots, \\ &= -i\sqrt{(\tfrac{1}{2}m)} && m=n+1 \\ &=0 && \text{otherwise.} \end{aligned}$$

m may be given by several variables taking integral values, or continuous ranges of values, or both; but n must be given by corresponding variables taking the same values; and the summation or integration giving the product of two matrices in the representation, which corresponds to the product of the dynamical variables in the algebra, must be specified or implied. We write

$$[a|\xi\eta|a]=[a|\xi|a][a|\eta|a] \tag{1.21}$$

and

$$(a_m|\xi\eta|a_o)=\textstyle\sum_n (a_m|\xi|a_n)(a_n|\eta|a_o), \tag{1.22}$$

with the understanding that $\sum_n$ is to be interpreted as required. We abbreviate the unit matrix $[a|1|a]$ to $[a|a]$. When there is no material ambiguity as to what representation is implied, $(a_m|\xi|a_n)$ may be abbreviated to $(m|\xi|n)$, and since m and n take the same values it is often convenient to replace them by n and n', especially when they are given explicitly by sets of variables; e.g., $(j,\ m|L|j',\ m')$. When convenient we designate the representation by the symbol of a dynamical variable or by those of a set of commuting dynamical variables, α, β, $\cdots$, say, represented by diagonal matrices for which the elements vanish when $n\neq m$. We may then use the diagonal elements, α', α'', $\cdots$; β', β'', $\cdots$; $\cdots$; of these matrices, the characteristic values of the dynamical variables, to designate the rows and columns of the representation, and write

$$[\alpha, \beta, \cdots|\xi|\alpha, \beta, \cdots]$$

for the matrix representing ξ in this representation, and

$$(\alpha', \beta', \cdots | \xi | \alpha'', \beta'', \cdots)$$

for an arbitrary element of that matrix. The use of quantum numbers combines these notations, dropping the primes when there is no ambiguity. E.g., if q and p are represented as above, and if

$$E = \tfrac{1}{2}(p^2/\mu + \mu\omega^2 q^2)$$

$E/\hbar\omega$ is represented by the matrix

$$\begin{matrix} \frac{1}{2} & 0 & 0 & 0 & \cdot\cdot \\ 0 & \frac{3}{2} & 0 & 0 & \cdot\cdot \\ 0 & 0 & \frac{5}{2} & 0 & \cdot\cdot \\ 0 & 0 & 0 & \frac{7}{2} & \cdot\cdot \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot \end{matrix}$$

and if we put $E = (n+\frac{1}{2})\hbar\omega$, we may write

$$(n-1 | q\sqrt{(2\mu\omega/\hbar)} | n) = i\sqrt{(n/2)}, \quad n = 0, 1, 2, \cdots$$
$$(n | q\sqrt{(2\mu\omega/\hbar)} | n-1) = -i\sqrt{(n/2)},$$

other elements vanishing (though the phases of the elements are not determined by

$$\tfrac{1}{2}(p^2/\mu + \mu\omega^2 q^2) = (n+\tfrac{1}{2})\hbar\omega).$$

If we have two representations, a and a^*, for which the variables designating the rows and columns take the same values, and if dynamical variables ξ, η, are represented in repesentation a by the matrices $[a|\xi|a]$, $[a|\eta|a]$, $\cdots$, we may define (different) dynamical variables ξ^*, η^*, $\cdots$, represented by the same matrices in the representation a^*, so that

$$\begin{matrix} (a_m^* | \xi^* | a_n^*) = (a_m | \xi | a_n), \\ (a_m^* | \eta^* | a_n^*) = (a_m | \eta | a_n) \\ \cdots \qquad\qquad \cdots \end{matrix} \tag{1.32}$$

or

$$\begin{matrix} [a^* | \xi^* | a^*] = [a | \xi | a], \\ [a^* | \eta^* | a^*] = [a | \eta | a] \\ \cdots \qquad\qquad \cdots \end{matrix} \tag{1.31}$$

The dynamical variables ξ^*, η^*, $\cdots$, then satisfy kinematical conditions of the same form as those satisfied by ξ, η, $\cdots$. Conversely, if we have two sets of dynamical variables ξ, η, $\cdots$, and ξ^*, η^*, $\cdots$, satisfying kinematical conditions of the same form, and if ξ, η, $\cdots$, are represented in a representation a by the matrices $[a|\xi|a]$, $[a|\eta|a]$, $\cdots$, we can define a representation a^* in which ξ^*, η^*, $\cdots$, are represented by the same matrices, so that

$$\begin{matrix} (a_m^* | \xi^* | a_n^*) = (a_m | \xi | a_n) \\ (a_m^* | \eta^* | a_n^*) = (a_m | \eta | a_n), \\ \cdots \qquad\qquad \cdots \end{matrix} \tag{1.32}$$

or

$$\begin{matrix} [a^* | \xi^* | a^*] = [a | \xi | a], \\ [a^* | \eta^* | a^*] = [a | \eta | a] \\ \cdots \qquad\qquad \cdots \end{matrix} \tag{1.31}$$

Such a symbol as $F(\xi, \eta, \cdots)$ will designate a function of the dynamical variables ξ, η, $\cdots$; $F(\xi, \eta, \cdots)$ will be itself a dynamical variable and may be defined in general by a matrix representing it in some representation of the dynamical variables ξ, η, $\cdots$, which matrix must be invariant for any change of representation that leaves the matrices representing the dynamical variables ξ, η, $\cdots$, invariant; the most usual functions can be regarded as convergent power series in the dynamical variables, the order of factors in the terms being, however, important.

If the dynamical variables ξ^*, η^*, $\cdots$, satisfy kinematical conditions of the same form as those satisfied by ξ, η, $\cdots$; and if $F(\xi, \eta, \cdots)$ denotes a function of the dynamical variables ξ, η, $\cdots$; then $F(\xi^*, \eta^*, \cdots)$ will denote the same function of the dynamical variables ξ^*, η^*, $\cdots$; e.g., if

$$F(\xi, \eta, \cdots) = a\xi + b\eta + \cdots,$$
$$F(\xi^*, \eta^*, \cdots) = a\xi^* + b\eta^* + \cdots.$$

If we require the function of ξ^*, η^*, $\cdots$, that is equal to $F(\xi, \eta, \cdots)$, we shall denote it by $F^*(\xi^*, \eta^*, \cdots)$ so that we shall have

$$F^*(\xi^*, \eta^*, \cdots) = F(\xi, \eta, \cdots) \tag{1.4}$$

and these may represent the same physical quantity, such as energy, electric moment, etc. E.g.,

$$\left.\begin{matrix} \text{if } q = q^* \cos\epsilon + p^* \sin\epsilon \\ p = -q^* \sin\epsilon + p^* \cos\epsilon \end{matrix}\right\} (\text{so that } qp - pq = i\hbar$$

and $q^*p^* - p^*q^* = i\hbar$ follow from each other), and if $F(q, p) = a(q^2 - p^2) + b(qp + pq)$, then

$$F(q, p) = F^*(q^*, p^*) = (a\cos 2\epsilon - b\sin 2\epsilon)(q^{*2} + p^{*2}) + (a\sin 2\epsilon + b\cos 2\epsilon)(q^*p^* + p^*q^*).$$

These notations are believed to be in current use;[9] except that it is not usual to write $[a|\xi|a]$ for the matrix representing the dynamical vari-

[9] P. A. M. Dirac, reference 6, Chapter 3, and p. 109; E. C. Kemble, reference 6, p. 268.

able ξ in the representation a, distinguished on the one hand from its elements and on the other from matrices representing ξ in other representations; and that the idea of a function of dynamical variables is often confined to a function of a set of commuting dynamical variables.

2. CHANGE OF REPRESENTATION AND TRANSFORMATION OF DYNAMICAL VARIABLES[10]

We change from one representation a with rows and columns denoted by m, n, to another representation b with rows and columns denoted by u, v, by using transformation matrices $[b|a]$, $[a|b]$, such that for any dynamical variable ξ,

$$\begin{aligned}[b|\xi|b]&=[b|a][a|\xi|a][a|b],\\ [a|\xi|a]&=[a|b][b|\xi|b][b|a].\end{aligned} \tag{2.01}$$

$[b|a)$ and $[a|b]$ must be such that $(b_u|a_m)$ and $(a_m|b_u)$ are complex conjugates and that

$$\begin{aligned}[b|a][a|b]&=[b|b],\\ [a|b][b|a]&=[a|a].\end{aligned} \tag{2.02}$$

In general there need be no correlation implied between the values of m or n and the values of u or v. If a one-to-one correspondence is given between these values so that the methods of summation or integration correspond, we may conveniently denote them by the same symbol m or n, and denote the representations by a and a^*; and we shall call the change of representation a canonical change of representation. (This specialization is not usually made explicitly.) The transformation matrix is then unitary and can be expressed in the form

$$[a|a^*]=e^{i[s]} \tag{2.03}$$

so that

$$[a^*|a]=e^{-i[s]}$$

where $[s]$ is a Hermitian matrix with elements $(m|s|n)$ such that $(m|s|n)$ and $(n|s|m)$ are complex conjugates.

We shall regard

$$[a^*|\xi|a^*]=e^{-i[s]}[a|\xi|a]e^{i[s]} \tag{2.04}$$

or

$$[a|\xi|a]=e^{i[s]}[a^*|\xi|a^*]e^{-i[s]}$$

as the standard form for a canonical change of representation.

[10] P. A. M. Dirac, reference 6, Chapter 3; E. C. Kemble, reference 6, p. 394.

We transform from one set of dynamical variables ξ, η, $\cdots$, to another set α, β, $\cdots$, by defining each set as functions of the other. In general the new set will satisfy kinematical conditions of different forms from those satisfied by the old. If they satisfy kinematical conditions of the same form, we may conveniently denote them by ξ^*, η^*, $\cdots$; and if ξ, η, $\cdots$ are represented by matrices in a representation a, then ξ^*, η^*, $\cdots$ will be represented by the same matrices in another representation a^*. The change from a to a^* is canonical and

$$\begin{aligned}[a|\xi^*|a]&=[a|a^*][a^*|\xi^*|a^*][a^*|a]\\ &=[a|a^*][a|\xi|a][a^*|a].\end{aligned} \tag{2.05}$$

Thus if U is a unitary function of ξ, η, $\cdots$, which has matrix elements in the representation a given by[11]

$$(a_m|U|a_n)=(a_m|a_n^*),$$

so that

$$[a|U|a]=[a|a^*], \tag{2.06}$$

we have

$$\begin{aligned}\xi^*&=U\xi U^{-1},\\ \eta^*&=U\eta U^{-1}\end{aligned} \tag{2.071}$$

and indeed, for any function $F(\xi, \eta, \cdots)$,

$$F(\xi^*, \eta^*, \cdots)=UF(\xi, \eta, \cdots)U^{-1} \tag{2.072}$$

(contrast this with

$$F^*(\xi^*, \eta^*, \cdots)=F(\xi, \eta, \cdots) \tag{1.4}$$

by definition). $(a_n|U^{-1}|a_m)$ and $(a_m|U|a_n)$ are complex conjugates and $UU^{-1}=U^{-1}U=1$. Such a transformation from dynamical variables ξ, η, $\cdots$, to dynamical variables ξ^*, η^*, $\cdots$, satisfying the same kinematical conditions is called a contact transformation of the dynamical variables.

U being unitary, it can be expressed in the form

$$U=e^{iS/\hbar}, \tag{2.08}$$

where S is a real function of the dynamical variables (i.e., the matrices representing it are Hermitian). If

$$[a|a^*]=e^{i[s]}, \tag{2.03}$$

then

$$(a_m|S|a_n)=\hbar(m|s|n)$$

or

$$[a|S|a]=\hbar[s]. \tag{2.09}$$

[11] P. A. M. Dirac, reference 6, p. 109.

We shall write

$$S=S(\xi,\eta,\cdots)=S^*(\xi^*,\eta^*,\cdots) \quad (2.10)$$

and shall regard the form

$$\xi^*=e^{iS/\hbar}\xi e^{-iS/\hbar},\quad \eta^*=e^{iS/\hbar}\eta e^{-iS/\hbar},\quad \cdots, \quad (2.11)$$

or

$$\xi=e^{-iS/\hbar}\xi^* e^{iS/\hbar},\quad \eta=e^{-iS/\hbar}\eta^* e^{iS/\hbar},\quad \cdots,$$

as the standard form for a contact transformation. A function of the original dynamical variables $\xi, \eta, \cdots$, is then transformed as follows:

$$F^*(\xi^*,\eta^*,\cdots)=F(\xi,\eta,\cdots) =e^{-iS/\hbar}F(\xi^*,\eta^*,\cdots)e^{iS/\hbar}. \quad (2.12)$$

E.g., if $qp-pq=i\hbar$, $S=cp$ gives

$$q^*=e^{icp/\hbar}qe^{-icp/\hbar}=q+c,$$
$$p^*=e^{icp/\hbar}pe^{-icp/\hbar}=p$$

and $F^*(p^*,q^*)=F(p,q)=F(p^*,q^*-c)$.[12]

$$[a^*|F|a^*]=e^{-i[s]}[a|F|a]e^{i[s]} \quad \text{(from (2.04))}$$

and

$$F^*(\xi^*,\eta^*,\cdots)=e^{-iS/\hbar}F(\xi^*,\eta^*,\cdots)e^{iS/\hbar} \quad (2.12)$$

are actually just different ways of writing the same result. The formula

$$e^{-(i/\hbar)K}Fe^{(i/\hbar)K}\equiv F+\frac{i}{\hbar}(FK-KF)$$
$$+\frac{1}{2}\left(\frac{i}{\hbar}\right)^2((FK-KF)K-K(FK-KF))+\cdots$$
$$\equiv F+[K,F]+\tfrac{1}{2}[K,[K,F]]+\cdots$$
$$+\frac{1}{r!}[K,[\cdots,[K,F]\cdots]]+\cdots, \quad (2.13)$$

where the Poisson bracket $[K, F]$ is defined by

$$[K,F]=-\frac{i}{\hbar}(KF-FK), \quad (2.14)$$

is useful for evaluating these expressions. If $F(\xi,\eta,\cdots)$ and $S^*(\xi^*,\eta^*,\cdots)$ are known as power series, this formula gives

$$F^*(\xi^*,\eta^*,\cdots)=F(\xi^*,\eta^*,\cdots) +[S(\xi^*,\eta^*,\cdots),F(\xi^*,\eta^*,\cdots)] +\tfrac{1}{2}[S(\xi^*,\eta^*,\cdots),[S(\xi^*,\eta^*,\cdots), \times F(\xi^*,\eta^*,\cdots)]]+\cdots. \quad (2.15)$$

3. THE MECHANICAL PROBLEM. PERTURBATION THEORY

We suppose that the Hamiltonian function H of a system, and the various physical properties in which we are interested, such as electric moment $\mathbf{e}$ and polarizability Φ, are given in terms of dynamical variables the kinematical relations between which are known. Our object is to find matrices representing H, $\mathbf{e}$, Φ, etc., in a representation b in which H has only diagonal elements: to find values for u, v, and for E_u, such that

$$(b_u|H|b_v)=E_u \quad v=u \qquad (3.1)$$
$$=0 \quad \text{otherwise},$$

and to find $(b_u|\mathbf{e}|b_v)$, $(b_u|\Phi|b_v)$, $\cdots$ etc. We divide this into a zero-order problem, and perturbation theory.[13]

The zero-order problem is to transform to dynamical variables $\xi, \eta, \cdots$, such as components of angular momentum and normal coordinates and momenta, in terms of which H has nearly a form H_0 for which we know such a representation, a. H is conveniently expressed as a power series in ϵ

$$H(\xi,\eta,\cdots)=H_0(\xi,\eta,\cdots)+\epsilon H_1(\xi,\eta,\cdots) +\epsilon^2H_2(\xi,\eta,\cdots)+\cdots, \quad (3.2)$$

where ϵ is small (e.g., $\epsilon=h/\nu I$ where ν is a vibration frequency, I a moment of inertia), and where values for m, n, and E_{0m} are known, so

$$(a_m|H_0|a_n)=E_{0m} \quad m=n \qquad (3.3)$$
$$=0 \quad \text{otherwise.}$$

We must also be able to find the matrix components in the representation a of functions of $\xi, \eta, \cdots$, in particular $(a_m|H_1|a_n)$, $(a_m|H_2|a_n)$, $\cdots$, $(a_m|\mathbf{e}|a_n)$, $(a_m|\Phi|a_n)$, $\cdots$. If we apply the canonical change of representation for which

$[a|a^*]=e^{i\epsilon[s]}$ to (from (2.03))

$$[a|H|a]=[a|H_0|a]+\epsilon[a|H_1|a]+\epsilon^2[a|H_2|a]+\cdots, \text{ (from (3.2))}$$

[12] P. A. M. Dirac, reference 6, p. 94.
[13] P. A. M. Dirac, reference 6, Chapter 8; E. C. Kemble, reference 6, p. 395; A. Ruark and H. C. Urey, reference 6, p. 596.

changing it to

$$[a^*|H|a^*]=[a^*|H_0^*|a^*]+\epsilon[a^*|H_1^*|a^*]+\epsilon^2[a^*|H_2^*|a^*]+\cdots$$

we obtain, expanding in powers of ϵ,[14]

$$[a^*|H_0^*|a^*]=[a|H_0|a], \tag{3.41}$$
$$[a^*|H_1^*|a^*]=[a|H_1|a]+i([a|H_0|a][s]-[s][a|H_0|a]), \tag{3.42}$$
$$\begin{aligned}[a^*|H_2^*|a^*]&=[a|H_2|a]+i([a|H_1|a][s]-[s][a|H_1|a])\\&\quad+\tfrac{1}{2}i^2(([a|H_0|a][s]-[s][a|H_0|a])[s]-[s]([a|H_0|a][s]-[s][a|H_0|a]))\\&=[a|H_2|a]+\tfrac{1}{2}i([a|H_1+H_1^*|a][s]-[s][a|H_1+H_1^*|a]),\end{aligned} \tag{3.43}$$

... ...[15]

If we apply the equivalent contact transformation from $\xi, \eta, \cdots$, to $\xi^*, \eta^*, \cdots$, given by $U=e^{is/\hbar}$ to (from (2.08)

$$H(\xi, \eta, \cdots)=H_0(\xi, \eta, \cdots)+\epsilon H_1(\xi, \eta, \cdots)+\epsilon^2 H_2(\xi, \eta, \cdots)+\cdots \text{ (from (3.2))}$$

transforming it to

$$H^*(\xi^*, \eta^*, \cdots)=F_0^*(\xi^*, \eta^*, \cdots)+\epsilon F_1^*(\xi^*, \eta^*, \cdots)+\epsilon^2 F_2^*(\xi^*, \eta^*, \cdots)+\cdots[16]$$

we obtain, expanding in powers of ϵ,

$$F_0^*(\xi^*, \eta^*, \cdots)=H_0(\xi^*, \eta^*, \cdots), \tag{3.51}$$
$$F_1^*(\xi^*, \eta^*, \cdots)=H_1(\xi^*, \eta^*, \cdots)+[S(\xi^*, \eta^*, \cdots), H_0(\xi^*, \eta^*, \cdots)], \tag{3.52}$$
$$\begin{aligned}F_2^*(\xi^*, \eta^*, \cdots)&=H_2(\xi^*, \eta^*, \cdots)+[S(\xi^*, \eta^*, \cdots), H_1(\xi^*, \eta^*, \cdots)]\\&\quad+\tfrac{1}{2}[S(\xi^*, \eta^*, \cdots), [S(\xi^*, \eta^*, \cdots), H_0(\xi^*, \eta^*, \cdots)]]\\&=H_2(\xi^*, \eta^*, \cdots)+\tfrac{1}{2}[S(\xi^*, \eta^*, \cdots), H_1(\xi^*, \eta^*, \cdots)+F_1^*(\xi^*, \eta^*, \cdots)].\end{aligned} \tag{3.53}$$

... ...

These results we shall write

$$H_0^*=H_0, \tag{3.51}$$
$$H_1^*=H_1+[S, H_0], \tag{3.52}$$
$$H_2^*=H_2+\tfrac{1}{2}[S, H_1+H_1^*], \tag{3.53}$$
$$H_3^*=H_3+[S, H_2]+\tfrac{1}{6}[S, [S, 2H_2+H_2^*]], \tag{3.54}$$

... ...

We remove terms from H by taking suitable forms for s or S. If the zero-order approximation is not degenerate, i.e., has no coincident energy values, we remove first terms of order ϵ; then terms up to ϵ^3, then to ϵ^7, and so on; only terms that are functions of H_0 remain, i.e., that are diagonal in the representation a^*, a^{**}, $\cdots$, that we have reached. If the zero-order approximation is degenerate, and if H_1, H_2, $\cdots$ contain terms that remove the degeneracy, these terms also remain, and we must solve a zero-order perturbation problem, changing the representation or transforming the dynamical variables suitably, to approach the representation b.

If this process converged we should approach the desired representation b, and could apply the same transformations to $\mathbf{e}$ and Φ. In any case, if ϵ is small, the first few transformations must give results with errors of order ϵ^2, ϵ^4, ϵ^8, $\cdots$, in the elements of $\mathbf{e}$ and Φ as well as in E_m.

Successive canonical changes of representation are especially useful when we want results for only a few values of m or n; contact transformations where the functions can be expanded in power series in the dynamical variables. We shall give tables of forms for S which give various values of $[S, H_0]$ for standard forms of H_0, and of the resulting higher order terms that are introduced into H.

[14] Born, Heisenberg, and Jordon, Zeits. f. Physik **35**, 565 (1925).

[15] Often applied: e.g., J. H. Van Vleck, Phys. Rev. **33**, 484 (1929); O. M. Jordahl, Phys. Rev. **43**, 87 (1934).

[16] Note that

$$H_0^*(\xi^*, \eta^*, \cdots)=H_0(\xi^*, \eta^*, \cdots)+\epsilon[S(\xi^*, \eta^*, \cdots), H_0(\xi^*, \eta^*, \cdots)]+\cdots\neq F_0^*(\xi^*, \eta^*, \cdots).$$

A Practical Method for the Solution of Certain Problems in Quantum Mechanics by Successive Removal of Terms from the Hamiltonian by Contact Transformations of the Dynamical Variables

Part II. Power Series in a Coordinate and Its Conjugate Momentum. The Anharmonic Oscillator by Perturbation Theory

L. H. THOMAS
Mendenhall Laboratory of Physics, The Ohio State University, Columbus, Ohio

(Received December 18, 1941)

The non-commutative algebra of polynomials in a coordinate and its conjugate momentum is reduced to common algebra by Weyl's method, and tables are given for facilitating its use. It is shown how the problem of the anharmonic oscillator can be solved by contact transformations, and tables are given for removing terms up to the third degree in its coordinate and momentum and for finding the modified Hamiltonian up to the fourth degree, which is as far as is ordinarily required in molecular theory. To illustrate the power of the method, the energy is computed up to terms of the eighth degree for energy containing terms to that degree (in the coordinate only), obtaining Dunham's result together with the constant terms of that order not given by him.

4. POWER SERIES IN A COORDINATE AND ITS CONJUGATE MOMENTUM

IF
$$qp-pq=i\hbar \tag{4.1}$$
and if we deal only with terms up to the fourth degree in these variables, we may apply non-commutative algebra directly to expressions such as pqp^2 and take explicit account of relations such as $p^3q=pqp^2-2i\hbar p^2$. It is, however, convenient even here, and almost essential for extension to terms of higher degree, to adopt a definite convention for a standard combination of such terms as p^3q, p^2qp, pqp^2, and qp^3, and to construct a general theorem for combining these combinations.

We define $\{p^rq^s\}$ by the expansion
$$e^{\alpha p+\beta q}\equiv\sum_r\sum_s \alpha^r\beta^s\frac{\{p^rq^s\}}{r!s!}, \tag{4.2}$$
where α and β are ordinary numbers, so that, for instance,
$$\{p^2q^2\}\equiv\tfrac{1}{6}(p^2q^2+pqpq+pq^2p+qpqp+q^2p^2+qp^2q)$$
and $\{p^rq^s\}$ is $r!s!/(r+s)!$ times the sum of the $(r+s)!/r!s!$ terms distinct in non-commutative algebra which become p^rq^s in ordinary algebra.[1]

We shall write
$$\{\sum_{r,s}a_{rs}p^rq^s\}\equiv\sum_{r,s}a_{rs}\{p^rq^s\}$$
for a linear combination of terms of non-commutative algebra constructed in this way. With the aid of 4.1 any polynomial of non-commutative algebra can be expressed in this form in one and only one way (Table I). E.g.,
$$p^3q=\{p^3q-\tfrac{3}{2}i\hbar p^2\}$$
$$=\tfrac{1}{6}(p^3q+p^2qp+pqp^2+qp^3)-\tfrac{3}{2}i\hbar p^2.$$

TABLE I. $\hbar=1$, $[q, p]=1$.

$1=\{1\}$	$q^4=\{q^4\}$
	$q^3p=\{q^3p+\frac{3}{2}iq^2\}$
$q=\{q\}$	$q^2pq=\{q^3p+\frac{1}{2}iq^2\}$
$p=\{p\}$	$qpq^2=\{q^3p-\frac{1}{2}iq^2\}$
	$pq^3=\{q^3p-\frac{3}{2}iq^2\}$
$q^2=\{q^2\}$	$q^2p^2=\{q^2p^2+2iqp-\frac{1}{2}\}$
$qp=\{qp+\frac{1}{2}i\}$	$qpqp=\{q^2p^2+iqp\}$
$pq=\{qp-\frac{1}{2}i\}$	$qp^2q=\{q^2p^2+\frac{1}{2}\}$
$p^2=\{p^2\}$	$pq^2p=\{q^2p^2+\frac{1}{2}\}$
	$pqpq=\{q^2p^2-iqp\}$
$q^3=\{q^3\}$	$p^2q^2=\{q^2p^2-2iqp-\frac{1}{2}\}$
$q^2p=\{q^2p+iq\}$	$qp^3=\{qp^3+\frac{3}{2}ip^2\}$
$qpq=\{q^2p\}$	$pqp^2=\{qp^3+\frac{1}{2}ip^2\}$
$pq^2=\{q^2p-iq\}$	$p^2qp=\{qp^3-\frac{1}{2}ip^2\}$
$qp^2=\{qp^2+ip\}$	$p^3q=\{qp^3-\frac{3}{2}ip^2\}$
$pqp=\{qp^2\}$	$p^4=\{p^4\}$
$p^2q=\{qp^2-ip\}$	
$p^3=\{p^3\}$	

Reintroduce $\hbar$ in terms of degree $1/qp$ of the leading term, $\hbar^2$ in terms of degree $1/q^2p^2$, and so on.

[1] This definition follows H. Weyl, Zeits. f. Physik **46**, 1 (1927). The definition given by M. Born and P. Jordan, Zeits. f. Physik **34**, 858 (1925) and expanded by Born, Jordan, and Heisenberg, Zeits. f. Physik **35**, 557 (1926), leads to considerably more complicated results. It should be noted that we use differentiation only of the polynomials of common algebra.

TABLE II. $\hbar=1$, $[q, p]=1$, $[\{A\}, \{B\}]=\{C\}$.

A \ B	q	p	q^2	qp	p^2	q^3	q^2p	qp^2	p^3
q	0	1	0	q	$2p$	0	q^2	$2qp$	$3p^2$
p	-1	0	$-2q$	$-p$	0	$-3q^2$	$-2qp$	$-p^2$	0
q^2	0	$2q$	0	$2q^2$	$4qp$	0	$2q^3$	$4q^2p$	$6qp^2$
qp	$-q$	p	$-2q^2$	0	$2p^2$	$-3q^3$	$-q^2p$	qp^2	$3p^3$
p^2	$-2p$	0	$-4qp$	$-2p^2$	0	$-6q^2p$	$-4qp^2$	$-2p^3$	0
q^3	0	$3q^2$	0	$3q^3$	$6q^2p$	0	$3q^4$	$6q^3p$	$9q^2p^2-\frac{3}{2}$
q^2p	$-q^2$	$2qp$	$-2q^3$	q^2p	$4qp^2$	$-3q^4$	0	$3q^2p^2+\frac{1}{2}$	$6qp^3$
qp^2	$-2qp$	p^2	$-4q^3p$	$-qp^2$	$2p^3$	$-6q^3p$	$-3q^2p^2-\frac{1}{2}$	0	$3p^4$
p^3	$-3p^2$	0	$-6qp^2$	$-3p^3$	0	$-9q^2p^2+\frac{3}{2}$	$-6qp^3$	$-3p^4$	0

Reintroduce $\hbar^2$ in terms of degree $1/q^2p^2$ of the leading term, $\hbar^4$ in terms of degree $1/q^4p^4$, and so on.

We make use of the lemma[2]

$$e^{\alpha p+\beta q}\equiv e^{\alpha p}e^{\beta q}e^{\frac{1}{2}\hbar\alpha\beta}\equiv e^{\beta q}e^{\alpha p}e^{-\frac{1}{2}\hbar\alpha\beta}, \quad (4.31)$$

or rather its corollary

$$e^{\alpha p+\beta q}e^{\gamma p+\delta q}=e^{(\alpha+\gamma)p}e^{(\beta+\delta)q}e^{\frac{1}{2}i\hbar(\beta\gamma-\alpha\delta)}, \quad (4.32)$$

to obtain the result, fundamental for working with polynomials in p and q, ($D_q(A)$ means $\partial/\partial q$ acting on A but not on B, and so on),

$$\begin{aligned}\{A\}\{B\}&\equiv\{e^{\frac{1}{2}i\hbar(D_q(A)D_p(B)-D_p(A)D_q(B))}AB\}\\&\equiv\{AB\}+\tfrac{1}{2}i\hbar\left\{\frac{\partial A}{\partial q}\frac{\partial B}{\partial p}-\frac{\partial A}{\partial p}\frac{\partial B}{\partial q}\right\}\\&\quad+\tfrac{1}{2}(\tfrac{1}{2}i\hbar)^2\left\{\frac{\partial^2A}{\partial q^2}\frac{\partial^2B}{\partial p^2}\right.\\&\quad\left.-2\frac{\partial^2A}{\partial p\partial q}\frac{\partial^2B}{\partial p\partial q}+\frac{\partial^2A}{\partial p^2}\frac{\partial^2B}{\partial q^2}\right\}\\&\quad+\tfrac{1}{6}(\tfrac{1}{2}i\hbar)^3\left\{\frac{\partial^3A}{\partial q^3}\frac{\partial^3B}{\partial p^3}\right.\\&\quad-3\frac{\partial^3A}{\partial q^2\partial p}\frac{\partial^3B}{\partial q\partial p^2}+3\frac{\partial^3A}{\partial q\partial p^2}\frac{\partial^3B}{\partial q^2\partial p}\\&\quad\left.-\frac{\partial^3A}{\partial p^3}\frac{\partial^3B}{\partial q^3}\right\}+\cdots. \quad (4.4)\end{aligned}$$

In particular the Poisson bracket

$$\begin{aligned}[\{A\},\{B\}]&=\left\{\frac{\partial A}{\partial q}\frac{\partial B}{\partial p}-\frac{\partial A}{\partial p}\frac{\partial B}{\partial q}\right\}\\&\quad+\tfrac{1}{6}(\tfrac{1}{2}i\hbar)^2\left\{\frac{\partial^3A}{\partial q^3}\frac{\partial^3B}{\partial p^3}-3\frac{\partial^3A}{\partial q^2\partial p}\frac{\partial^3B}{\partial q\partial p^2}\right.\\&\quad\left.+3\frac{\partial^3A}{\partial q\partial p^2}\frac{\partial^3B}{\partial q^2\partial p}-\frac{\partial^3A}{\partial p^3}\frac{\partial^3B}{\partial q^3}\right\}+\cdots. \quad (4.5)\end{aligned}$$

[2] H. Weyl, reference 1; the lemma follows from the properties of the group.

This last result connects quantum theory contact transformations with those of classical theory, enabling correction terms to be written down (Table II).

5. THE HARMONIC OSCILLATOR—EXCURSUS ON NOTATION

If
$$H=\frac{1}{2m}p^2+\frac{m\omega^2}{2}q^2, \quad (5.1)$$

where
$$qp-pq=i\hbar \quad (4.1)$$

the usual reduction of H to a diagonal matrix gives[3]

$$(n|H/k\omega|n)=n+\tfrac{1}{2}\quad n=0, 1, 2, \cdots \quad (5.21)$$

the remaining components vanishing,

$$\begin{aligned}&\left(n+1\left|\frac{1}{\sqrt{\hbar}}(p/\sqrt{(m\omega)}+i\sqrt{(m\omega)}q)\right|n\right)\\&\quad=\sqrt{(2(n+1))}\quad n=0, 1, 2, \cdots \quad (5.22)\end{aligned}$$

the remaining components vanishing,

$$\begin{aligned}&\left(n\left|\frac{1}{\sqrt{\hbar}}(p/\sqrt{(m\omega)}-i\sqrt{(m\omega)}q)\right|n+1\right)\\&\quad=\sqrt{(2(n+1))}\quad n=0, 1, 2, \cdots \quad (5.23)\end{aligned}$$

the remaining components vanishing.

[3] W. Heisenberg, Zeits. f. Physik **33**, 879 (1925); M. Born and P. Jordan, Zeits. f. Physik **34**, 858 (1925).

The derivation is:—if $H=\frac{1}{2}p^2+\frac{1}{2}q^2$, $qp-pq=i$; write $p+iq=\alpha$, $p-iq=\beta$, and $4H=\alpha\beta+\beta\alpha$, $\alpha\beta-\beta\alpha=2$; $H\alpha-\alpha H$ reduces to α; hence either $(n|\alpha|m)=0$ or $H_n-H_m=1$; and $(n|\alpha|m)(m|\beta|n)=|(n|\alpha|m)|^2=2H_n-1=2H_m+1$. The series $\cdots, H_m, H_m-1, H_m-2, \cdots$ must terminate since it must be positive, and can do so only for the value $\frac{1}{2}$; the values of H are $\cdots, \frac{5}{2}, \frac{3}{2}, \frac{1}{2}$; the representation can be chosen so that $(n+1|\alpha|n)$ and $(n|\beta|n+1)$ are real and positive, and the result follows.

These variables and matrices have been used for the examples of Part I of this paper.

Here we have used m for the mass of a particle moving under a restoring force proportional to the displacement such as to give a classical frequency of oscillation $\nu=\omega/2\pi$. This is customary in theoretical work. The use of $\hbar$ and ω removes most 2π's from the formulae.[4]

In applications to diatomic molecules it is usual to use μ for the reduced mass in place of m and to write $-kq$ for the restoring force so that we write

$$H=\frac{1}{2\mu}p^2+\frac{1}{2}kq^2. \tag{5.11}$$

It is usual to reserve ν for the actual wave number of a line (in cm^{-1}) rather than for the frequency (in rev./sec.), and to reserve ω for a value (in cm^{-1}) corresponding to the classical frequency for small oscillations.[5] If they occur in the same context we shall write ν_{osc} and ω_{osc} for the former, frequency of oscillation in rev./sec. and radians/sec., and $\tilde{\nu}$ and $\tilde{\omega}$ for the latter, observed wave number and corrected wave number, respectively.

For polyatomic molecules it is convenient to use dimensionless coordinates, $s=\sqrt{(m\omega_{osc}/\hbar)}q$ with the corresponding momenta

$$p_s=\sqrt{(\hbar/m\omega_{osc})}p$$

so that

$$H=\frac{1}{2}\hbar\omega_{osc}(p_s^2/\hbar^2+s^2)=\frac{1}{2}\frac{\tilde{\omega}h}{c}(p_s^2/\hbar^2+s^2) \tag{5.12}$$

for the harmonic approximation for a single normal coordinate.[6]

For the purposes of this paper it is convenient to choose the unit of mass so that $m=1$, that of time so that $\omega=1$, and that of length so that $\hbar=1$, (this has already been done in Tables I and II). Thus we work with

$$H=\tfrac{1}{2}p^2+\tfrac{1}{2}q^2, \tag{5.13}$$

$$qp-pq=i. \tag{5.14}$$

To return to the usual units, replace p by $p/\sqrt{(m\hbar\omega)}$, q by $q\sqrt{(m\omega/\hbar)}$, H by $H/\hbar\omega$.

We can obtain the matrix components of $\{p^uq^v\}$ in the representation given by 5.21, 5.22, 5.23, as follows: From 5.22, 5.23,

$$(n+r|\,(p+iq)^r(p-iq)^s\,|n+s)=\frac{\sqrt{((n+r)!(n+s)!)}}{n!}2^{r+s/2}$$

n, r, s, positive integers or zero, (5.3)

the remaining components vanishing. From

$$e^{\alpha(p+iq)+\beta(p-iq)}=e^{\alpha\beta}e^{\alpha(p+iq)}e^{\beta(p-iq)} \tag{4.31}$$

we obtain

$$(n+r|\,\{(p+iq)^r(p-iq)^s\}\,|n+s)$$

$$=\frac{\sqrt{((n+r)!(n+s)!)}}{n!}2^{r+s/2}\left(1+\frac{r\cdot s}{1\cdot(n+1)}\left(\frac{1}{2}\right)+\frac{r(r-1)s(s-1)}{1\cdot2\cdot(n+1)(n+2)}\left(\frac{1}{2}\right)^2+\cdots\right)$$

$$=\frac{\sqrt{((n+r)!(n+s)!)}}{n!}2^{r+s/2}F(-r,\,-s;\,n+1;\,\tfrac{1}{2}) \tag{5.4}$$

$n+r$, $n+s\geqslant 0$, the remaining components vanishing (if n is negative, the first $-n$ terms of the series

[4] P. A. M. Dirac, *Quantum Mechanics*, second edition (Oxford, 1934), pp. 90, 133; H. Weyl, *The Theory of Groups and Quantum Mechanics* (Methuen, 1931), p. 51.
[5] G. Herzberg, *Molecular Spectra and Molecular Structure* I (Prentice Hall, 1939), Chapter 3.
[6] S. Silver and W. H. Shaffer, J. Chem. Phys. **9**, 599 (1941).

vanish); which, since

$$\{p^u q^v\} \equiv \sum_{s=0}^{u+v} \left\{ \frac{(p+iq)^{u+v-s}(p-iq)^s}{2^{u+v}} \right\} \sum_b \frac{u!v!(-1)^b}{b!(r-s+b)!(s-b)!(v-b)!}$$

$$\equiv \sum_{s=0}^{u+v} \left\{ \frac{(p+iq)^{u+v-s}(p-iq)^s}{2^{u+v}} \right\} \frac{u!}{s!(r-s)!} F(-s, -v; r-s+1; -1) \tag{5.5}$$

enables us to write down the matrix components of $\{p^u q^v\}$, e.g., returning to the usual units,

$$(n+4|\,\{p^3q^2\}\,|n+1) = \sqrt{(m\omega)}\left(\frac{\hbar}{2}\right)^{5/2} \tfrac{1}{2}(n+3)\sqrt{((n+4)(n+3)(n+2))}, \quad (n+2) \geqslant 0.$$

6. THE ANHARMONIC OSCILLATOR

If
$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}q^2 + \epsilon H_1(p, q), \tag{6.1}$$
where
$$qp - pq = i \tag{5.14}$$

we may remove a term $c\{p^r q^s\}$ from H_1 by a contact transformation to p^*, q^*, with a function $\epsilon c\{S\}$ given by

$$[\tfrac{1}{2}p^{*2} + \tfrac{1}{2}q^{*2}, \{S\}] = \{p^{*r} q^{*s}\} \tag{6.21}$$

i.e., (4.5)

$$q^* \frac{\partial S}{\partial p^*} - p^* \frac{\partial S}{\partial q^*} = p^{*r} q^{*s} \tag{6.22}$$

so that the equation to find S is for the anharmonic oscillator of exactly the same form as in classical dynamics. (In returning to the usual units, replace S by $S/\hbar$). We then have (3.53)

$$H^* = H = \tfrac{1}{2}p^{*2} + \tfrac{1}{2}q^{*2} + \epsilon(H_1(p^*, q^*) + [c\{S\}, \tfrac{1}{2}p^{*2} + \tfrac{1}{2}q^{*2}]) + \epsilon^2[c\{S\}, H_1(p^*, q^*)$$

$$+ \tfrac{1}{2}[c\{S\}, \tfrac{1}{2}p^{*2} + \tfrac{1}{2}q^{*2}]] + \frac{\epsilon^3}{2}[c\{S\}, [c\{S\}, H_1(p^*, q^*) + \tfrac{1}{3}[c\{S\}, \tfrac{1}{2}p^{*2} + \tfrac{1}{2}q^{*2}]]] + \cdots. \tag{6.3}$$

To save writing it is convenient to drop the asterisks, remembering that both $\{S\}$ and the transformed H are in terms of the transformed variables.

Thus for s odd we may take

$$S = \frac{p^{r+1}q^{s-1}}{(r+1)} + \frac{(s-1)}{(r+1)(r+3)} p^{r+3}q^{s-3} + \cdots + \frac{(s-1)\cdots 2}{(r+1)(r+3)\cdots(r+s)} p^{r+s} \tag{6.41}$$

and for r odd

$$S = -\frac{p^{r-1}q^{s+1}}{(s+1)} - \frac{(r-1)}{(s+1)(s+3)} p^{r-3}q^{s+3} - \cdots - \frac{(r-1)\cdots 2}{(s+1)(s+3)\cdots(s+r)} q^{s+r}. \tag{6.42}$$

If r and s are both odd, these two expressions differ by a multiple of $(p^2+q^2)^{r+s/2}$, and the sum of either and any multiple of this may be taken. (The transformation function reducing H to a function of $\tfrac{1}{2}p^{*2} + \tfrac{1}{2}q^{*2}$ is itself in fact only determined up to an additive function of $\tfrac{1}{2}p^{*2} + \tfrac{1}{2}q^{*2}$, though the reduced form of H is completely determined.)

If r and s are both even, $c\{p^r q^s\}$ cannot be removed, but can be reduced to $C\{(p^2+q^2)^{r+s/2}\}$. If $\backsim$ means can be reduced to,

TABLE III.

$$
\begin{array}{ll}
\{q^2\} \backsim \tfrac{1}{2}\{(p^2+q^2)\} & = \tfrac{1}{2}(p^2+q^2) \\
\{p^2\} \backsim \tfrac{1}{2}\{(p^2+q^2)\} & = \tfrac{1}{2}(p^2+q^2) \\
 & \\
\{q^4\} \backsim \tfrac{3}{8}\{(p^2+q^2)^2\} & = \tfrac{3}{8}((p^2+q^2)^2+1) \\
\{q^2p^2\} \backsim \tfrac{1}{8}\{(p^2+q^2)^2\} & = \tfrac{1}{8}((p^2+q^2)^2+1) \\
\{p^4\} \backsim \tfrac{3}{8}\{(p^2+q^2)^2\} & = \tfrac{3}{8}((p^2+q^2)^2+1) \\
 & \\
\{q^6\} \backsim \tfrac{5}{16}\{(p^2+q^2)^3\} & = \tfrac{5}{16}((p^2+q^2)^3+5(p^2+q^2)) \\
\{q^4p^2\} \backsim \tfrac{1}{16}\{(p^2+q^2)^3\} & = \tfrac{1}{16}((p^2+q^2)^3+5(p^2+q^2)) \\
\{q^2p^4\} \backsim \tfrac{1}{16}\{(p^2+q^2)^3\} & = \tfrac{1}{16}((p^2+q^2)^3+5(p^2+q^2)) \\
\{p^6\} \backsim \tfrac{5}{16}\{(p^2+q^2)^3\} & = \tfrac{5}{16}((p^2+q^2)^3+5(p^2+q^2)) \\
 & \\
\{q^8\} \backsim \tfrac{35}{128}\{(p^2+q^2)^4\} & = \tfrac{35}{128}((p^2+q^2)^4+14(p^2+q^2)^3+9) \\
\{q^6p^2\} \backsim \tfrac{5}{128}\{(p^2+q^2)^4\} & = \tfrac{5}{128}((p^2+q^2)^4+14(p^2+q^2)^2+9) \\
\{q^4p^4\} \backsim \tfrac{3}{128}\{(p^2+q^2)^4\} & = \tfrac{3}{128}((p^2+q^2)^4+14(p^2+q^2)^2+9) \\
\{q^2p^6\} \backsim \tfrac{5}{128}\{(p^2+q^2)^4\} & = \tfrac{5}{128}((p^2+q^2)^6+14(p^2+q^2)^2+9) \\
\{p^8\} \backsim \tfrac{35}{128}\{(p^2+q^2)^4\} & = \tfrac{35}{128}((p^2+q^3)^4+14(p^2+q^2)^2+9)
\end{array}
$$

Replace p by $p/\sqrt{(m\hbar\omega)}$, q by $q\sqrt{(m\omega/\hbar)}$.

TABLE IV. $\hbar=1$, $[q, p]=1$, $[\frac{1}{2}(p^2+q^2), \{S\}]=\{H_1\}$

H_1	S	H_1	S
$1-1$	0	$q^4-\frac{3}{8}(p^2+q^2)^2$	$\frac{5}{8}q^3p+\frac{3}{8}qp^3$
		q^3p	$-\frac{5}{32}q^4+\frac{3}{16}q^2p^2+\frac{3}{32}p^4$
q	p	$q^2p^2-\frac{1}{8}(p^2+q^2)^2$	$-\frac{1}{8}q^3p+\frac{1}{8}qp^3$
p	$-q$	qp^3	$-\frac{3}{32}q^4-\frac{3}{16}q^2p^2+\frac{5}{32}p^4$
		$p^4-\frac{3}{8}(p^2+q^2)^2$	$-\frac{3}{8}q^3p-\frac{5}{8}qp^3$
$q^2-\frac{1}{2}(p^2+q^2)$	$\frac{1}{2}qp$		
qp	$-\frac{1}{4}q^2+\frac{1}{4}p^2$	q^5	$q^4p+\frac{4}{3}q^2p^3+\frac{8}{15}p^5$
$p^2-\frac{1}{2}(p^2+q^2)$	$-\frac{1}{2}qp$	q^4p	$-\frac{1}{5}q^5$
		q^3p^2	$\frac{1}{3}q^2p^3+\frac{2}{15}p^5$
q^3	$q^2p+\frac{2}{3}p^3$	q^2p^3	$-\frac{2}{15}q^5-\frac{1}{3}q^3p^2$
q^2p	$-\frac{1}{3}q^3$	qp^4	$\frac{1}{5}p^5$
qp^2	$\frac{1}{3}p^3$	p^5	$-\frac{8}{15}q^5-\frac{4}{3}q^3p^2-qp^4$
p^3	$-\frac{2}{3}q^3-qp^2$		

Replace p by $p/\sqrt{(m\hbar\omega)}$, q by $q\sqrt{(m\omega/\hbar)}$, H by $H/\hbar\omega$, S by $S/\hbar$.

$$\{q^{2n}\} \backsim \frac{2n-1}{1}\{q^{2n-2}p^2\} \backsim \frac{(2n-1)(2n-3)}{1\cdot 3}\{q^{2n-4}p^4\} \backsim \cdots \backsim \{p^{2n}\} \backsim [\{(q^2+p^2)^n\} \Big/$$

$$1+\frac{n}{1}\frac{1}{(2n-1)}+\frac{n(n-1)}{1\cdot 2}\frac{1\cdot 3}{(2n-1)(2n-3)}+\cdots+1] = \{(q^2+p^2)^n\}\frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n} \qquad (6.5)$$

(Tables III and IV).

For the anharmonic oscillator the difference from classical mechanics enters in the Poisson brackets, (e.g., in the usual units, $[\{q^4p\}, \{qp^4\}] = \{15q^4p^4-6\hbar q^2p^2-\frac{3}{2}\hbar^2\}$ of which the first term only would occur in classical mechanics).

When H_1 can be expressed as a power series in p and q, a series for S is obtained reducing H to a power series in $\{(p^2+q^2)^n\}$ together with terms of higher order than ϵ; a second transformation reduces H to a power series in $\{(p^2+q^2)^n\}$ together with terms of higher order than ϵ^3, and so on.

$(n|H|n)$ can then be found from 5.3, or reduction formulae giving $\{(p^2+q^2)^n\}$ in terms of (p^2+q^2) can be built up (Table III). If terms beyond ϵ^2 can be neglected and if the terms of order ϵ in H are of order three or lower in p and q, $(n|H|n)$ can be written down immediately from Table V.

7. THE SLIGHTLY ANHARMONIC OSCILLATOR

If
$$H=\left(\frac{1}{2m}p^2+\frac{1}{2}m\omega^2q^2\right)+\epsilon c_3q^3+\epsilon^2c_4q^4+\cdots, \qquad (7.1)$$

TABLE V. $H_1=H_1'+H_1''$, $[\frac{1}{2}(p^2+q^2), \{S'\}]=H_1'-H_1'^*$, $[\frac{1}{2}(p^2+q^2), H_1'^*]=0$, $\frac{1}{2}[\{S'\}, H_1''+H_1''^*]=C \backsim D$

H_1' \ H_1''	$\{q\}$	$\{p\}$	$\{q^2\}$	$\{qp\}$	$\{p^2\}$	$\{q^3\}$	$\{q^2p\}$	$\{qp^2\}$	$\{p^3\}$
$\{q\}$	$\begin{matrix}\{-\frac{1}{2}\}\\ -\frac{1}{2}\end{matrix}$	$\begin{matrix}0\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}p\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{3}{4}q^2\}\\ -\frac{3}{4}(p^2+q^2)\end{matrix}$	$\begin{matrix}\{-qp\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}p^2\}\\ -\frac{1}{4}(p^2+q^2)\end{matrix}$	$\begin{matrix}0\\0\end{matrix}$
$\{p\}$	$\begin{matrix}0\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}\}\\ -\frac{1}{2}\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}p\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}p\}\\0\end{matrix}$	$\begin{matrix}0\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^2\}\\ -\frac{1}{4}(p^2+q^2)\end{matrix}$	$\begin{matrix}\{-qp\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{3}{2}p^2\}\\ -\frac{1}{4}(p^2+q^2)\end{matrix}$
$\{q^2\}$	$\begin{matrix}\{-\frac{1}{4}q\}\\0\end{matrix}$	$\begin{matrix}\{\frac{1}{4}p\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^2+\frac{1}{2}p^2\}\\0\end{matrix}$	$\begin{matrix}0\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^2+\frac{1}{2}p^2\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{3}{4}q^3\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{4}q^2p\}\\0\end{matrix}$	$\begin{matrix}\{\frac{1}{4}qp^2\}\\0\end{matrix}$	$\begin{matrix}\{\frac{1}{4}p^3\}\\0\end{matrix}$
$\{qp\}$	$\begin{matrix}\{-\frac{1}{4}p\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{4}q\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}qp\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{4}q^2-\frac{1}{4}p^2\}\\ -\frac{1}{4}(p^2+q^2)\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}qp\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{3}{4}q^2p\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{4}q^3-\frac{1}{2}qp^2\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^2p-\frac{1}{4}p^3\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{3}{4}qp^2\}\\0\end{matrix}$
$\{p^2\}$	$\begin{matrix}\{\frac{1}{4}q\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{4}p\}\\0\end{matrix}$	$\begin{matrix}\{\frac{1}{2}q^2-\frac{1}{2}p^2\}\\0\end{matrix}$	$\begin{matrix}0\\0\end{matrix}$	$\begin{matrix}\{\frac{1}{2}q^2-\frac{1}{2}p^2\}\\0\end{matrix}$	$\begin{matrix}\{\frac{3}{4}q^3\}\\0\end{matrix}$	$\begin{matrix}\{\frac{1}{4}q^2p\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{4}qp^2\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{3}{4}p^3\}\\0\end{matrix}$
$\{r^3\}$	$\begin{matrix}\{-\frac{1}{2}q^2-p^2\}\\ -\frac{3}{4}(p^2+q^2)\end{matrix}$	$\begin{matrix}\{qp\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^3\}\\0\end{matrix}$	$\begin{matrix}\{\frac{1}{2}q^2p-p^3\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^3\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{3}{2}q^4-3q^2p^2+\frac{1}{2}\}\\ -1\frac{5}{16}(p^2+q^2)^2-\frac{7}{16}\end{matrix}$	$\begin{matrix}\{-2qp^3\}\\0\end{matrix}$	$\begin{matrix}\{\frac{3}{2}q^2p^2+\frac{1}{4}-p^4\}\\ -3\frac{1}{16}(p^2+q^2)^2+\frac{5}{16}\end{matrix}$	$\begin{matrix}\{3qp^3\}\\0\end{matrix}$
$\{q^2p\}$	$\begin{matrix}0\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^2\}\\ -\frac{1}{4}(p^2+q^2)\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^2p\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^3\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^2p\}\\0\end{matrix}$	$\begin{matrix}0\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}q^4\}\\ -3\frac{1}{16}(p^2+q^2)^2-\frac{3}{16}\end{matrix}$	$\begin{matrix}\{-q^2p\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{3}{2}q^2p^2+\frac{1}{4}\}\\ -3\frac{1}{16}(p^2+q^2)^2+1\frac{1}{16}\end{matrix}$
$\{qp^2\}$	$\begin{matrix}\{-\frac{1}{2}p^2\}\\ -\frac{1}{4}(p^2+q^2)\end{matrix}$	$\begin{matrix}0\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}qp^2\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}p^3\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}qp^2\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{3}{2}q^2p^2+\frac{1}{4}\}\\ -3\frac{1}{16}(p^2+q^2)^2+1\frac{1}{16}\end{matrix}$	$\begin{matrix}\{-qp^3\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}p^4\}\\ -3\frac{1}{16}(p^2+q^2)^2-\frac{3}{16}\end{matrix}$	$\begin{matrix}0\\0\end{matrix}$
$\{r^3\}$	$\begin{matrix}\{qp\}\\0\end{matrix}$	$\begin{matrix}\{-q^2-\frac{1}{2}p^2\}\\ -\frac{3}{4}(p^2+q^2)\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}p^3\}\\0\end{matrix}$	$\begin{matrix}\{-q^3+\frac{1}{2}qp^2\}\\0\end{matrix}$	$\begin{matrix}\{-\frac{1}{2}p^3\}\\0\end{matrix}$	$\begin{matrix}\{\frac{1}{2}q^2p\}\\0\end{matrix}$	$\begin{matrix}\{-q^4+\frac{3}{2}q^2p^2+\frac{1}{4}\}\\ -3\frac{1}{16}(p^2+q^2)^2+\frac{5}{16}\end{matrix}$	$\begin{matrix}\{-2q^3p\}\\0\end{matrix}$	$\begin{matrix}\{-3q^2p^2+\frac{1}{2}q-\frac{3}{2}p^4\}\\ -3\frac{1}{16}(p^2+q^2)^2-\frac{3}{16}\end{matrix}$

Replace p by $p/\sqrt{(m\hbar\omega)}$, q by $q/\sqrt{(m\omega/\hbar)}$, H by $H/\hbar\omega$, S by $S/\hbar$.

where

$$qp-pq=i\hbar \tag{4.1}$$

transformation with

$$S=\epsilon c_3\left\{\frac{p^*q^{*2}}{m\omega^2}+\frac{2}{3}\frac{p^{*3}}{m^3\omega^4}\right\} \qquad \text{(Table IV)} \quad (7.2)$$

gives to order ϵ^2

$$H^*=\left\{\frac{1}{2m}p^{*2}+\frac{1}{2}m\omega^2q^{*2}+\epsilon^2\left(\left(c_4-\frac{3}{2}\frac{c_3{}^2}{m\omega^2}\right)q^{*4}-\frac{3c_3{}^2}{m^3\omega^4}p^{*2}q^{*2}+\frac{\hbar^2c_3{}^2}{2m^3\omega^4}\right)+\cdots\right\}, \quad \text{(Table V)} \quad (7.3)$$

which a further transformation reduces to

$$H^{**}=\hbar\omega\left(\frac{1}{2m\hbar\omega}p^{**2}+\frac{1}{2}\frac{m\omega}{\hbar}q^{**2}\right)$$

$$+\epsilon^2\left(\left(\frac{3}{8}\frac{c_4}{m^2\omega^2}-\frac{15}{16}\frac{c_3{}^2}{m^3\omega^4}\right)\left(\left(\frac{1}{m\omega}p^{**2}+m\omega q^{**2}\right)^2+\hbar^2\right)+\frac{\hbar^2c_3{}^2}{2m^3\omega^4}\right) \quad \text{(Tables III, V)} \quad (7.4)$$

giving for the nth state, since

$$\frac{1}{2m\hbar\omega}p^{**2}+\frac{1}{2}\frac{m\omega}{\hbar}q^{**2}=n+\tfrac{1}{2},$$

$$E_n=(n+\tfrac{1}{2})\hbar\omega+\frac{\epsilon^2\hbar^2}{m^2\omega^2}\left(\tfrac{3}{2}c_4(n^2+n+\tfrac{1}{2})-\frac{c_3{}^2}{m\omega^2}\left(\frac{30n^2+30n+11}{8}\right)\right)+\cdots \tag{7.5}$$

Further, if

$$\mathbf{e}=\mathbf{e}_0+\epsilon\mathbf{e}_1q+\epsilon^2\mathbf{e}_2q^2+\cdots, \tag{7.6}$$

transformation with 7.2 gives

$$\mathbf{e}=\mathbf{e}^*(p^*,q^*)=\mathbf{e}_0+\epsilon\mathbf{e}_1q^*+\epsilon^2\left(\mathbf{e}_2q^{*2}-c_3\mathbf{e}_1\left(\frac{q^{*2}}{m\omega^2}+2\frac{p^{*2}}{m^3\omega^4}\right)\right)+\cdots \quad \text{(Table II).} \quad (7.7)$$

To this order the further transformation makes no change, and the matrix components of **e** in a system in which H is diagonal can be written down at once.[7]

8. DUNHAM'S RESULTS

To obtain the terms of the next two orders in E for the anharmonic oscillator, Dunham's[8] results, Table II (formula 4.5) were extended to cover $[\{A\}, \{B\}] = \{C\}$ up to terms of order q^8 in C. If

$$H = \tfrac{1}{2}p^2 + \tfrac{1}{2}q^2 + aq^3 + bq^4 + cq^5 + dq^6 + eq^7 + fq^8 + \cdots, \tag{8.01}$$

where

$$qp - pq = i, \tag{8.02}$$

transformation with

$$S = a\{pq^2 + \tfrac{2}{3}p^3\} \quad \text{(Table IV)} \tag{8.03}$$

removes aq^3 giving by 3.51, 3.52, $\cdots$ and using Table II repeatedly,

$$\begin{aligned} H = \{&\tfrac{1}{2}(q^2+p^2) + bq^4 + \tfrac{1}{2}a^2(-3q^4-6q^2p^2+1) + cq^5 + ab(-4q^5-8q^3p^2+4q) \\ &+ \tfrac{1}{3}a^3(12q^5+12q^3p^2-18q+24qp^4) + dq^6 + ac(-5q^6-10q^4p^2+10q^2) + \tfrac{1}{2}a^2b(20q^6+32q^4p^2 \\ &-56q^2+48q^2p^4-16p^2) + \tfrac{1}{8}a^4(-60q^6-108q^4p^2+156q^2+96q^2p^4+120p^2-48p^6) + eq^7 \\ &+ ad(-6q^7-12q^5p^2+20q^3) + \tfrac{1}{2}a^2c(30q^7+60q^5p^2-140q^3+80q^3p^4-80qp^2) \\ &+ \tfrac{1}{6}a^3b(-120q^7-240q^5p^2+576q^3+32q^3p^4+576qp^2-192qp^6) + \tfrac{1}{30}a^5(360q^7+720q^5p^2 \\ &-1728q^3+1440q^3p^4-960qp^6) + fq^8 + ae(-7q^8-14q^6p^2+35q^4) + \tfrac{1}{2}a^2d(42q^8+96q^6p^2 \\ &-300q^4+120q^4p^4-240q^2p^2+20) + \tfrac{1}{6}a^3c(-210q^8-480q^6p^2+1620q^4-200q^4p^4+1920q^2p^2 \\ &-180-480q^2p^5+240p^4) + \tfrac{1}{24}a^4b(840q^8+1920q^6p^2-6528q^4+2560q^4p^4-3840q^2p^2+864 \\ &-2304q^2p^6-2560p^4+384p^8) + \tfrac{1}{144}a^6(-2520q^8-5760q^6p^2+19584q^4+30528q^2p^2-1728 \\ &-19200q^2p^6-5760p^4+1920p^8) + \cdots\}. \end{aligned} \tag{8.04}$$

A second transformation with

$$\begin{aligned} S = \{&b(\tfrac{5}{8}q^3p + \tfrac{3}{8}qp^3) + \tfrac{1}{2}a^2(-\tfrac{9}{8}q^3p - \tfrac{15}{8}qp^3) + c(q^4p + \tfrac{4}{3}q^2p^3 + \tfrac{8}{15}p^5) \\ &+ ab(-4q^4p - 8q^2p^3 - 16\tfrac{2}{5}p^5 + 4p) + \tfrac{1}{3}a^3(12q^4p + 20q^2p^3 + 6\tfrac{4}{5}p^5 - 18p)\} \quad \text{(Table IV)} \end{aligned} \tag{8.05}$$

reduces

$$\{bq^4 + \tfrac{1}{2}a^2(-3q^4-6q^2p^2+1) + cq^5 + ab(-4q^5-8q^3p^2+4q) + \tfrac{1}{3}a^3(12q^5+12q^3p^2-18q+24qp^4)\}$$

to

$$\{\tfrac{3}{8}b(q^2+p^2)^2 + \tfrac{1}{16}a^2(-15(q^2+p^2)^2+8)\},$$

giving, by 3.51, 3.52, $\cdots$, and using Table II repeatedly,

$$\begin{aligned} H = \{&\tfrac{1}{2}(q^2+p^2) + \tfrac{3}{8}b(q^2+p^2)^2 + \tfrac{1}{16}a^2(-15(q^2+p^2)^2+8) + dq^6 + ac(-5q^6-10q^4p^2+10q^2) \\ &+ \tfrac{1}{32}a^2b(467q^6+713q^4p^2-1274q^2+777q^2p^4-58p^2-45p^6) + \tfrac{1}{128}a^4(-1311q^6-2781q^4p^2 \\ &+3900q^2+771q^2p^4+840p^2-543p^6) + \tfrac{1}{32}b^2(-55q^6-69q^4p^2+72q^2+27q^2p^4-36p^2+9p^6) + eq^7 \\ &+ ad(-6q^7-12q^5p^2+20q^3) + a^2c\tfrac{1}{32}(681q^7+1341q^5p^2-3245q^3+1664q^3p^4-1856qp^2 \\ &+256qp^6) + a^3b\tfrac{1}{24}(-1497q^7-4161q^5p^2+8247q^3-3296q^3p^4+9210qp^2-2028qp^6) \end{aligned}$$

[7] M. Born and P. Jordan, Zeits. f. Physik **34**, 858 (1925).
[8] J. L. Dunham, Phys. Rev. **41**, 713, 721 (1932).

$$+a^5{}_{1}{}_{8}{}_{0}(2970q^7+7020q^5p^2-18318q^3+11430q^3p^4-19890qp^2+3460qp^6)$$
$$+bc{}_{1}{}_{4}{}_{8}(-207q^7-555q^5p^2+507q^3-256q^3p^4+768qp^2)+ab^2{}_{4}(69q^7+243q^5p^2-307q^3$$
$$+194q^3p^4-396qp^2)+fq^8+ae(-7q^8-14q^6p^2+35q^4)+{}_{4}bd(-15q^8-27q^6p^2+45q^4)$$
$$+{}_{6}c^2(-15q^8-60q^6p^2+60q^4-40q^4p^4+240q^2p^2-12)+{}_{8}a^2d(195q^8+519q^6p^2-1425q^4$$
$$+480q^4p^4-960q^2p^2+80)+{}_{1}{}_{2}abc(465q^8+1215q^6p^2-3051q^5+1570q^4p^4-3858q^2p^2$$
$$+288+384q^2p^6-448p^4)+{}_{1}{}_{2}{}_{8}b^3(425q^8+620q^6p^2-2310q^4+930q^4p^4+612q^2p^2+108q^2p^4$$
$$-342p^4+27p^8)+{}_{2}{}_{4}a^3c(-1725q^8-5595q^6p^2+13959q^4-3850q^4p^4+16794q^2p^2-1488$$
$$-1664q^2p^6+2688p^4-256p^8)+{}_{2}{}_{5}{}_{6}a^2b^2(-23125q^8-60116q^6p^2+185450q^4-65150q^4p^4$$
$$+2080740q^2p^2-20480-55380q^3p^6+45242p^4-405p^8)+{}_{3}{}_{0}{}_{7}{}_{2}a^4b(586614q^8+1927128q^6p^2$$
$$-5241036q^4+1897540q^2p^4-6127800q^2p^2+700416+246744q^2p^6-1905004p^4+216438p^8)$$
$$+{}_{1}{}_{0}{}_{2}{}_{4}a^6(-89903q^8-339260q^6p^2+841278q^4-259930q^4p^4+1210284q^2p^2-131072$$
$$-59324q^2p^6+358574p^4-42543p^8)+\cdots\}. \quad (8.06)$$

A further transformation removes the terms of the order of eq^7 and gq^9 and reduces H to the form (Table III)

$$H={}_{2}(q^2+p^2)+{}_{8}b(3(q^2+p^2)^2+3)+{}_{1}{}_{6}a^2(-15(q^2+p^2)^2-7)+{}_{1}{}_{6}d(5(q^2+p^2)^3+25(q^2+p^2))$$
$$+{}_{1}{}_{6}ac(-35(q^2+p^2)^3-95(q^2+p^2))+{}_{3}{}_{2}a^2b(225(q^2+p^2)^3+459(q^3+p^2))$$
$$+{}_{1}{}_{2}{}_{8}a^4(-705(q^2+p^2)^3-1155(q^2+p^2))+{}_{3}{}_{2}b^2(-17(q^2+p^2)^3-67(q^2+p^2))$$
$$+{}_{1}{}_{2}{}_{8}f(35(q^2+p^2)^4+490(q^2+p^2)^2+315)+{}_{1}{}_{2}{}_{8}ae(-315(q^2+p^2)^4-2730(q^2+p^2)^2-1155)$$
$$+{}_{1}{}_{2}{}_{8}bd(-165(q^2+p^2)^4-1770(q^2+p^3)^3-945)+{}_{2}{}_{5}{}_{6}c^2(-315(q^2+p^2)^4-2170(q^2+p^2)^2$$
$$-1107)+{}_{2}{}_{5}{}_{6}a^2d(2715(q^2+p^2)^4+17070(q^2+p^2)^2+6055)+{}_{1}{}_{2}{}_{8}abc(2415(q^2+p^2)^4$$
$$+14670(q^2+p^2)^2+5667)+{}_{2}{}_{5}{}_{6}b^3(6000(q^2+p^2)^4+13656(q^2+p^2)^2+1539)$$
$$+{}_{2}{}_{5}{}_{6}a^3c(-9765(q^2+p^2)^4-47730(q^2+p^2)^2-14777)+{}_{5}{}_{1}{}_{2}a^2b^2(-24945(q^2+p^2)^4$$
$$-248052(q^2+p^2)^2-40261)+{}_{1}{}_{0}{}_{2}{}_{4}a^4b(116325(q^2+p^2)^4+479970(q^2+p^2)^2+131817)$$
$$+{}_{2}{}_{0}{}_{4}{}_{8}a^6(-115755(q^2+p^2)^4-418110(q^2+p^2)^2-101479)+\text{terms of order } a^8. \quad (8.07)$$

This result agrees with that of Dunham by the W. K. B. method;[8] the terms ${}_{1}{}_{2}{}_{8}f\,315$ $-{}_{1}{}_{2}{}_{8}\,ae\,1155-{}_{1}{}_{2}{}_{8}\,bd\,945-\cdots-{}_{2}{}_{0}{}_{4}{}_{8}\,a^6\,101479$, not given by Dunham, check with an independent calculation by Dunham's method. The method used here may be regarded as differing from that of Dunham essentially only by expanding the potential energy in a power series first, and then expanding the transformation in powers of $\hbar$, while the W. K. B. method expands in powers of $\hbar$ first and Dunham expands the potential energy later. Except for the terms not given by Dunham, his method obtains the form for H more briefly than this method; but these terms, the only ones of order $\hbar^2$, require in his method nearly as much time as all the rest; in this method they are obtained with the others. We see also that, while for perturbation theory we require only the infinitesimal classical contact transformations, if we tried to determine the finite contact transformation corresponding to an arbitrary potential energy $V(q)$ we should be led back to the phase-integral and for the quantum theory corrections to the W. K. B. method.

Statistical Distribution Laws for Rate Processes

BRUCE LONGTIN
Illinois Institute of Technology, Chicago, Illinois
(Received May 4, 1942)

A re-examination of the fundamental hypotheses of statistical mechanics shows that the principle of the most probable distribution may be interpreted in such a way as to be applicable as well to systems which are not at equilibrium. The general method of deriving distribution laws for systems not at equilibrium is outlined. The use of this method is illustrated by deriving a general distribution law for systems in which the velocities are required to be non-Maxwellian. The distribution of velocities of molecules in any group having a specified average velocity is Maxwellian relative to the group average velocity. The average total energy of such a system is shown to be the sum of the kinetic energies of mass motion of the several groups of molecules plus the Maxwellian average energy corresponding to the given volume and temperature.

GROSS physical properties of a system of molecules may be calculated from the molecular constants either by kinetic theory methods or by those of statistical mechanics. The kinetic theory method as applied by Enskog,[1] Chapman,[2,3] and others consists in solving the Maxwell-Boltzmann integro-differential equation by successive approximations. The zeroth approximation is Maxwell's law[4] of velocity distribution. The second and third approximations have been obtained only with difficulty,[3] and little work has been done on higher approximations. The method at present is further limited to gases at low pressures.

The method of statistical mechanics, whether based on the canonical ensemble of Gibbs[5] or the principle of maximum probability[6] is more direct, giving the form of the statistical distribution function free of approximations. However, this method has been developed only for systems at equilibrium. The limitation to systems at equilibrium is not necessary, and may be removed by re-examining the fundamental hypotheses.

In developing a statistical mechanics which is applicable to rate processes, it is desirable to insure at the same time that it will apply as well to condensed systems. For this reason, the method of Jeans,[6] in which the element of the statistical ensemble is the whole system of molecules, will be followed. The language used is that of classical mechanics; however, only minor changes are required by the substitution of relativistic and wave mechanics.

GENERALIZED STATISTICAL MECHANICS

Definitions

In establishing a law connecting the gross physical properties of a system, a series of connected measurements are made on the same system. The procedure is then repeated on the same or another practically identical system under comparable conditions. The number of repetitions is increased until the experimenter is confident that further repetitions will not materially alter the average values of the measured properties.

The requirement of *comparable conditions* implies that certain gross parameters $g_1, g_2, \cdots$ of the system are the same in each repetition of the series of measurements. This specification is inadequate, as evidenced by the different values of the properties $x, y, \cdots$ obtained in the several sets of measurements. A complete specification would describe the system as existing in the state i whose properties are precisely $x_i, y_i, \cdots$. The inadequate specification leaves the system to assume any one of the detailed states $1, 2, \cdots, i, \cdots$ by chance.

The system may happen to assume the i state in a number nw_i out of n repetitions. This defini-

[1] D. Enskog, *Kinetische Theorie der Vorgänge in mässig verdünnten Gasen* (Dissertation, Upsala, 1917).
[2] S. Chapman, Phil. Trans. Roy. Soc. **A211**, 433 (1912); **216**, 279 (1916); **217**, 115 (1917).
[3] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge University Press, 1939).
[4] J. C. Maxwell, Phil. Trans. Roy. Soc. **157**, 49 (1867).
[5] J. Willard Gibbs, *The Collected Works* (Longmans, Green and Company, New York, 1938), Vol. 2.
[6] J. H. Jeans, Phil. Mag. [6] **5**, 597 (1903).

tion of the relative frequencies w_i is such that

$$\sum_i w_i = 1, \tag{1}$$

while the average value of the property x_i for this particular set of repetitions is

$$\bar{x} = \sum_i x_i w_i. \tag{2}$$

If the property x is constant over the states i, then $x_i = x_j = \cdots = \bar{x}$; the average value is the same as that for each state. In particular, $\sum_i w_i$ is the average value, $\bar{1}$, of unity.

As the number n of repetitions is increased, w_i may approach a limit denoted by W_i if the limit exists. Then since x_i depends only on the definition of the state i, the average value $\bar{x}$ will approach a limit X given by Eq. 2 when W_i is substituted for w_i. Let the operator D denote the value of the quantity operated on minus its limiting value. Thus $Dw_i = w_i - W_i$, $D\bar{x} = \bar{x} - X$, etc. By applying this operator to Eqs. 1 and 2, one finds

$$\sum_i Dw_i = D1 = 0, \tag{3}$$

and

$$D\bar{x} = \sum_i x_i Dw_i. \tag{4}$$

In the limit of large n, the operator D must be interpreted as denoting the change in the value operated on due to a trial variation in the values of the W_i's.

Fundamental Hypothesis

In order to give any meaning to statistical methods, it is necessary to make some assumption which in effect requires that

the set of limiting frequencies W_i exists and is determined by the conditions imposed on the set of repeated experiments.

The relation between the values W_i and the experimental conditions constitutes the statistical distribution law by means of which the limiting average values X, Y, $\cdots$ of various physical properties may be calculated.

Let a_i denote the *a priori* probability that in a single experiment the system will assume the particular detailed state i. In general this quantity is to be interpreted as the quantum weight of the state i, although other interpretations may be necessary in some cases. Further let $\mathfrak{W}_n(w_1, w_2, \cdots)$ denote the probability that out of a series of n experiments the state 1 was assumed exactly nw_1 times, the state 2 in nw_2 times, etc. If the performance of n repeated experiments constitutes a series of unrelated events, then

$$\mathfrak{W}_n(w_1, w_2 \cdots) = n! \prod_i (a_i^{nw_i}/nw_i!). \tag{5}$$

Using Stirling's approximation, if the number n of experiments is very large, Eq. 5 may be written in the form

$$\log \mathfrak{W}_n = n(\log a/w)_{Av}, \tag{6}$$

where $(\log a/w)_{Av}$ represents

$$\sum_i w_i \log a_i/w_i.$$

The hypothesis of a limiting distribution implies that in the limit of a very large number of experiments, the distribution $w_1 = W_1$, $w_2 = W_2$, $\cdots$ becomes very much (i.e., infinitely) more probable than any other distribution. For this to be true it is necessary that $(\log a/w)_{Av}$ have a value for this distribution which is at least infinitesimally greater than that for any other distribution. If $(\log a/w)_{Av}$ is a continuous function of the w_i's, then its value for the limiting distribution is a true maximum, and its variation

$$\sum_i (\log a_i/W_i - 1) Dw_i = 0 \tag{7}$$

for all small values of the Dw_i's which are consistent with the experimental conditions. Actually, $(\log a/w)_{Av}$ is defined only for integral values of the numbers nw_i, and may be considered a continuous function only in the limit of very large values of n.

Equation 7 is analogous to the usual principle of maximum probability,[6,7] but must be interpreted somewhat differently. In the usual treatment the distribution law concerns the number of independent molecules out of a very large total number which are to be found in a particular quantum state at equilibrium. The distribution implied by Eq. 7 is not concerned with numbers of molecules, but rather with the number of experiments out of a large number n, in which the system as a whole is observed to as-

[7] See, for example, E. H. Kennard, *Kinetic Theory of Gases* (McGraw-Hill Book Company, New York, 1938), Section 199.

sume a particular detailed state. The statistics of independent molecules may be considered as a special case in which the experimental system consists of only a single molecule. Furthermore, no assumption has been made which in any way limits the result to equilibrium states.

Distribution Law

To obtain the distribution law appropriate to any particular set of experimental conditions, Eq. 7 is to be solved subject to these conditions. The form of the resulting distribution law will depend on the nature of the restrictions which the system must satisfy.

In obtaining the solution, it is only necessary to consider a set of independent restrictions. If one of the restrictions implies a second, only the first need be included; the second will necessarily be satisfied. By using this principle extensively, calculation of the distribution law is greatly simplified.

The major features of the resulting distribution law are determined by equations which contain either the w_i's or the W_i's explicitly. Of these Eq. 7 is the only one which has been found to be generally true for systems at equilibrium. Any other such restriction involving the w_i's or W_i's explicitly can therefore be valid only under special conditions which are imposed by the experimenter.

The arbitrary restrictions imposed by the experimenter may often be expressed in the form

$$\sum_i x_i Dw_i = 0, \tag{8}$$

where the property x_i depends only on the state of the system and not explicitly on the frequencies w_i or W_i. For example, the rejection of data for which the average value $\bar{x}$ does not tend toward a definite limit leaves a set of data for which $D\bar{x}=0$ in the limit; Eq. 8 is satisfied. This is simply the requirement that the property x_i have a statistically reproducible value.

As a working hypothesis it will be assumed that no arbitrary restriction containing the w_i's or W_i's explicitly need be considered which cannot be expressed in the form of Eq. 8. In all cases so far investigated, of the various possible ways of expressing a particular experimental condition, the one most closely representing the physical facts has been in the form of Eq. 8. Since in practice such arbitrary restrictions can be executed only by rejecting data until the number of experiments falling in each detailed state has the specified value, it would also appear that any form other than Eq. 8 implies an unduly biased selection of data. The possible occurrence of such restrictions will therefore be left to be treated as special cases.

The remaining restrictions do not involve the W_i's or w_i's explicitly. Hence the required solution of Eq. 7 may be obtained by first eliminating the Dw_i's from an independent set of restrictions of the form of Eq. 8, together with Eq. 7. The resulting equation is then to be solved simultaneously with the remaining restrictions.

The result obtained in the first step by the method of Lagrange multipliers[6, 8, 9] is

$$\log W_i/a_i = \lambda_x x_i + \lambda_y y_i + \cdots + \lambda_1, \tag{9}$$

in which each coefficient λ must have the same value for all states of the system, but is otherwise as yet unrestricted. These coefficients may therefore prove to be functions of the parameters g_j whose values are independent of the detailed states of the system. The term λ_1 enters in order to insure that Eq. 3 can be satisfied. Each of the other terms corresponds to one of the independent equations, there being one term for each equation.

Of the remaining restrictions, those which contain $x_i, y_i \cdots$ explicitly impose certain conditions on these variables. The others, which do not contain these variables, can enter only through the λ's. It is usually convenient to leave the distribution law in the form of Eq. 9, remembering that additional restrictions remain to be satisfied.

Systems at Equilibrium

When the only experimental restriction is that the total energy E_i of the system have a statistically reproducible value (i.e., $D\bar{E}=0$ in the limit), then Eq. 9 reduces to

$$\log W_i/a_i = \lambda_E E_i + \lambda_1. \tag{10}$$

Equation 10 represents the Gibbs[5] canonical en-

[8] See, for example, R. B. Lindsay, *Introduction to Physical Statistics* (John Wiley and Sons, New York, 1941), p. 54.
[9] J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley and Sons, New York, 1940), Chapters 3 to 5.

semble, in which the coefficient λ_E has the experimental value[9] $-1/kT$, while λ_1 has the value A/kT, where A is the Helmholtz free energy, k the Boltzmann constant, and T the thermodynamic temperature.

In applying Eq. 10 it is assumed that the number of molecules N_{Bi} of each component B has a definite value. If instead it is only assumed that the value is statistically reproducible, then $D\bar{N}_B=0$ in the limit for each component. In this case, Eq. 10 also contains a term $\lambda_B N_{Bi}$ for each component so restricted, and then represents the grand canonical ensemble[5,10] in which each λ_B is $(1/kT)$ times the partial molal free energy (i.e., chemical potential) of the component. Other restrictions lead to other important forms of the ensemble.[10]

NON-MAXWELLIAN VELOCITY DISTRIBUTIONS

The most general problem of velocity distributions is that in which it is required that each molecule l have a different statistically reproducible average momentum $\bar{\mathbf{p}}_l$; the restriction $D\bar{\mathbf{p}}_l=0$ is satisfied by each molecule in the limit. Since these restrictions are independent of the requirement $D\bar{E}=0$, Eq. 10 must now contain an additional term $\boldsymbol{\lambda}_l\cdot\mathbf{p}_l$ for each molecule. By the classical method it may be shown that the coefficient λ_E still has the value $1/kT$. Expressing E_i as the sum of a potential energy V_i and the kinetic energy, the distribution law is found to be

$$\log W_i/a_i=-\frac{1}{kT}(V_i+\sum_l \mathbf{p}_{li}^2/2m_l)+\sum_l \boldsymbol{\lambda}_l\cdot\mathbf{p}_{li}+\lambda_1, \quad (11)$$

where m_l is the mass of molecule l.

By writing $\boldsymbol{\mu}_l$ in place of $m_l kT\boldsymbol{\lambda}_l$, Eq. 11 may be expressed in the form

$$\log W_i/a_i=-V_i/kT-\sum_l(\mathbf{p}_{li}-\boldsymbol{\mu}_l)^2/2m_l kT+(\lambda_1-\sum_l \boldsymbol{\mu}_l^2/2m_l kT). \quad (12)$$

The last term contains only parameters which are independent of the states i, and hence is a new arbitrary constant

$$\lambda_0=\lambda_1-\sum_l \boldsymbol{\mu}_l^2/2m_l kT. \quad (13)$$

[10] R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics* (University Press, Cambridge, 1939), Section 611.

It is thus seen that $\log W_i/a_i$ is the same function of the $(\mathbf{p}_{li}-\boldsymbol{\mu}_l)$'s and λ_0 for the non-Maxwellian distribution velocities that it would be of the $\mathbf{p}_{li}$'s and λ_1 for a Maxwellian distribution at the same temperature.

The summation over states of any function f_i which contains the momenta and the constants $\boldsymbol{\mu}_l$ only in the differences $(\mathbf{p}_{li}-\boldsymbol{\mu}_l)$ will be invariant to a change of the $\boldsymbol{\mu}_l$'s, provided only that the array of detailed states i is always transformed into itself by any shift of the origin of momenta in phase space. This requirement is satisfied *a fortiori* when the spectrum of permissible velocities forms an unlimited continuum.

In particular, from Eq. 1 it is seen that

$$\exp(-\lambda_0)=\sum_i a_i \times\exp[(-1/kT)(V_i+\sum_l(\mathbf{p}_{li}-\boldsymbol{\mu}_l)^2/2m_l)], \quad (14)$$

which is such a summation. Since $\exp(-\lambda_0)$ must therefore be invariant to a change in the $\boldsymbol{\mu}_l$'s, its value can be obtained by setting all of the $\boldsymbol{\mu}_l$'s equal to zero. It is then seen that λ_0 is equal to A_0/kT, where A_0 is the value of the Helmholtz free energy for the system at the given temperature and in the specified volume, for a Maxwellian distribution of velocities.

Similarly, the value of $(\mathbf{p}_{li}-\boldsymbol{\mu}_l)_{Av}$ must be invariant to a change of the $\boldsymbol{\mu}_l$'s and hence equal to its Maxwellian value. This value is zero since $(\mathbf{p}_{li}-\boldsymbol{\mu}_l)_{Av}$ is then the Maxwellian average momentum. In general $(\mathbf{p}_{li}-\boldsymbol{\mu}_l)_{Av}$ may be expanded as $\bar{\mathbf{p}}_{li}-\boldsymbol{\mu}_l$, since $\boldsymbol{\mu}_l$ is constant over the states i. Hence it has been shown that

$$\bar{\mathbf{p}}_{li}=\boldsymbol{\mu}_l; \quad (15)$$

the constants $\boldsymbol{\mu}_l$ must be equal to the specified average momentum.

The distribution law may now be expressed in the complete form

$$\log W_i/a_i=\frac{1}{kT}[A_0-V_i-\sum_l(\mathbf{p}_{li}-\bar{\mathbf{p}}_l)^2/2m_l]. \quad (16)$$

It is interesting to note that essentially Eq. 16 was obtained by Maxwell,[3,4] but for some reason the constant $\bar{\mathbf{p}}_l$ has always been interpreted as

$m_l\mathbf{v}_0$, where $\mathbf{v}_0$ is the velocity of a system of reference coordinates. For this reason the terms in $\bar{\mathbf{p}}_l$ have always been dropped. From the present derivation it is obvious that such a narrow interpretation was unwarranted.

In any particular velocity distribution problem the requirements will probably take the form that the total momentum of molecules of a particular group should have a statistically reproducible value. A detailed treatment of this case shows that the appropriate distribution law is obtained from Eq. 16 by setting each $\bar{\mathbf{p}}_l$ equal to the group average molecular momentum. Thus Eq. 16 represents the general distribution law for all problems of non-Maxwellian distribution of momenta, provided the array of detailed states i possesses the requisite transformation property.

Several important features of this distribution law may be observed. First it is noted that the distribution of velocities of molecules about the specified group average velocity is Maxwellian. In particular, if all molecules in the system are to have the same group average velocity, the law represents a mass of molecules moving with this velocity whose distribution of velocities relative to the center of gravity of the mass is Maxwellian. Furthermore, in general the distribution law is invariant to any change in velocity of the reference coordinates.

One application of this distribution law is illustrated by calculating the average energy of a system in which the velocity distribution is non-Maxwellian. Following the same procedure used to evaluate the $\mathfrak{y}_l$'s, one may define a quantity $[V_i+\sum_l (\mathbf{p}_{li}-\bar{\mathbf{p}}_l)^2/2m_l]_{Av}$ which is invariant to a change in the $\mathbf{p}_l$'s, and whose value is $\bar{E}_0$, the Maxwellian average total energy of the system for the same temperature and volume. By expanding the separate terms of this average, the result,

$$[\bar{V}_i+\sum_l (\mathbf{p}_{li}^2/2m_l)]_{Av} - \sum_l (\mathbf{p}_{li}\cdot\bar{\mathbf{p}}_l/m_l)_{Av}+\sum_l (\bar{\mathbf{p}}_l^2/2m_l)_{Av}=\bar{E}_0 \quad (17)$$

is obtained. The first term is the required average energy $\bar{E}$ while the second has the value $-\sum_l \bar{\mathbf{p}}_l^2/m_l$, and the third the value $\sum_l \bar{\mathbf{p}}_l^2/2m_l$. Making these substitutions, the average energy $\bar{E}$ is found to have the value

$$\bar{E}=\bar{E}_0+\sum_l \bar{\mathbf{p}}_l^2/2m_l. \quad (18)$$

Equation 18 expresses the average total energy of the system as the sum of the Maxwellian total energy, which consists of potential and thermal energy, plus the kinetic energy of mass motion of groups of molecules. The correctness of this result may be seen from considering the special case of, for example, a planetary system, in which the total energy of the system consists of the Maxwellian total energies of the individual planets plus the kinetic energy of mass motion of each planet.

This result is to be compared with the second approximation in the method of Enskog and Chapman,[3,7] which gives $W_i^0(1+\sum_l \mathbf{b}_l\cdot\mathbf{p}_{li})$ for the distribution function. Here W_i^0 is the Maxwellian distribution function, and $\mathbf{b}_l$ is a small arbitrary constant. Although this form agrees with Eq. 16 to this degree of approximation, it requires $\bar{E}=\bar{E}_0$ which cannot be generally true. It is only with the inclusion of the third and higher approximations that the expression for the average energy begins to take the proper form.

CONCLUSIONS

The method of deriving statistical distribution laws by solving Eq. 7 subject to the expressed independent experimental restrictions is a simple extension of the method used for systems at equilibrium. It is, however, in no way limited to equilibrium conditions. While the application of this method has been illustrated only by the derivation of a distribution law for systems in which the velocities are required to be non-Maxwellian, it may also be used to solve other statistical problems encountered in the study of rate processes.

It is of course recognized that the true test of this as well as of any theory is whether it agrees satisfactorily with experimental results, particularly when extended into new fields. At present this test has been applied only qualitatively, but with promising results. Consequently the author has no illusions that the distribution law for rate processes has been found in its ultimate form, but only hopes that a step has been taken in the right direction.

Statistical Distribution Laws for Rate Processes

2. Non-Uniform Distributions

BRUCE LONGTIN
Illinois Institute of Technology, Chicago, Illinois
(Received May 4, 1942)

The methods developed in the previous paper are used to determine the most general form of statistical distribution law suitable for the treatment of rate processes. The result is particularly applicable to relatively dense systems composed of soft molecules. It is based on the assumptions that each variable whose rate of change is to be measured must be statistically reproducible at all times, and that the molecular motions are not seriously discontinuous. Various special forms of the distribution law result from different ways of expressing the rate variables and experimental conditions. One form is derived for the case in which the rates of change of the average position and kinetic energies of the molecules are to be reproducible. Qualitatively, it agrees with the known frequent behavior of rate experiments, showing frictional dissipation of energy, and the tendency to return to equilibrium. Furthermore, it reduces to the classical distribution law under equilibrium conditions.

IN the previous paper[1] a general method of deriving statistical distribution laws was developed, and was applied to the case of a non-Maxwellian distribution of velocities. In this application, time was not considered explicitly as a variable. In the general study of rate processes, specific attention must be given to the way in which the experimental system changes with time. In so doing, the meaning of a statistical distribution and the form of the resulting distribution law must be somewhat modified.

DISTRIBUTION OF RATE EXPERIMENTS

The statistical method in general is concerned with the distribution of certain statistical elements among several categories. In the study of a rate process, the statistical element is a rate experiment on a single system, in which the significant properties x, y, $\cdots$ of the system changed in a particular way with time. The categories among which a series of rate experiments may be distributed are the different possible ways in which the set of properties x, y, $\cdots$ of the experimental system may have changed. Thus the statistical treatment of rate processes is concerned with the distribution of a series of repeated rate experiments among various detailed behaviors.

When a particular experiment is said to belong to the detailed behavior i it is implied that the properties x, y, $\cdots$ of the system have changed with time in the way ascribed to this detailed behavior. If the properties of the system have changed in some other way, then the particular rate experiment belongs to a different detailed behavior.

In particular, the variation of the property x with time for the detailed behavior i may be represented by a Taylor's series

$$x_i=\sum_m (\mathrm{d}^m x_i/\mathrm{d}t^m)_0(t-t_0)^m/m!, \qquad (1)$$

at least up to the time at which the first discontinuity appears in any one of the derivatives. The quantities $(\mathrm{d}^m x_i/\mathrm{d}t^m)_0$ are values of the various time derivatives of x_i appropriate to the reference time t_0. Any change in the numerical values of any of these quantities will cause x_i to change in a different way with time. Thus a particular rate experiment belongs to the detailed behavior i if in that experiment all of the quantities $(\mathrm{d}^m x/\mathrm{d}t^m)_0$, $(\mathrm{d}^m y/\mathrm{d}t^m)_0$, $\cdots$ have the values appropriate to this detailed behavior. If any one of them is different from the value required of the detailed behavior i, then the experiment belongs to another detailed behavior.

In classical mechanics, if the coordinates and momenta of all elementary particles in a system are given specified values at the time t_0, the behavior of the system is completely determined. These coordinates and momenta represent only a limited number of the $(\mathrm{d}^m x/\mathrm{d}t^m)_0$'s; the rest

[1] B. Longtin, J. Chem. Phys. **10**, 546 (1942).

are not independent variables and hence need not be specified separately. In quantum mechanics a somewhat different situation prevails.

After all of the n repeated rate experiments have been completed and the data recorded, a statistical analysis is in order. At this time the number nw_i of rate experiments belonging to each detailed behavior i is counted. It is at once evident that the value of nw_i is not in any way determined by the length of time elapsed during any portion of an individual experiment, nor does it depend on time in any other way. Nor does the total number n of experiments included in the statistical analysis depend on time. Thus each of the distribution frequencies w_i must be independent of time. Likewise the limit W_i approached by the frequency w_i as the number n is increased[2] must be independent of time. This constancy of the number of experiments belonging to a particular detailed behavior will be recognized as an analogue of Liouville's theorem.[3,4]

Average Values

The average value $\bar{x}$ of any property x_i is defined as

$$\bar{x}=\sum_i x_i w_i. \tag{2}$$

If the Taylor series expression given in Eq. 1 is valid, then

$$\bar{x}=\sum_i \sum_{m=0}^{\infty} (\mathrm{d}^m x_i/\mathrm{d}t^m)_0 w_i (t-t_0)^m/m!$$

$$=\sum_{m=0}^{\infty} (t-t_0)^m/m! \sum_i (\mathrm{d}^m x_i/\mathrm{d}t^m)_0 w_i. \tag{3}$$

Since neither the w_i's nor the $(\mathrm{d}^m x_i/\mathrm{d}t^m)$'s are functions of time, Eq. 3 expresses $\bar{x}$ as a function of time by a Taylor series. Such an expression is valid only up to the first discontinuity. A more general form must be sought when the motion of the system is seriously discontinuous.

One of the important requirements in a series of rate experiments is that each of the properties x, y, $\cdots$ whose rate of change is being studied must have a statistically reproducible value at all times during the progress of the experiment. This implies[1] that at all times

$$D\bar{x}=0 \tag{4}$$

in the limit as n is increased, where D denotes the change in the quantity operated on, due to passage from the limiting distribution W_1, W_2, $\cdots$ to some other distribution w_1, w_2, $\cdots$.

If the Taylor series representation is used,

$$D\bar{x}=\sum_m (t-t_0)^m/m! \sum_i (\mathrm{d}^m x_i/\mathrm{d}t^m)_0 Dw_i, \tag{5}$$

since only w_i is involved in a variation of the distribution. If $D\bar{x}$ is to vanish at all times, the coefficients of each $(t-t_0)^m/m!$ must be zero;

$$\sum_i (\mathrm{d}^m x_i/\mathrm{d}t^m)_0 Dw_i=0. \tag{6}$$

If Eq. 6 is satisfied for all integral values of m from zero up, then the quantity x will certainly be statistically reproducible at all times up to the first discontinuity.

Distribution Law

A statistical method is meaningless if the limiting frequencies W_i do not exist. In order that they shall exist, it is necessary[1] that

$$D(\log a/w)_{\mathrm{Av}}=0 \tag{7}$$

in the limit, $(\log a/w)_{\mathrm{Av}}$ being the average value of $\log a_i/w_i$, while a_i is the *a priori* probability that the experimental system will exhibit the detailed behavior i. The distribution law appropriate to any particular case is to be found by solving Eq. 7 subject to the experimental restrictions, and the requirement that the average value of 1 be unity.

In the case of rate experiments these restrictions include Eqs. 6. If none of the additional restrictions involves either the w_i's or W_i's explicitly, the method of solution used in the previous paper[1] leads to the distribution law

$$\log W_i/a_i=\lambda+\sum_{m=0}^{\infty}[\lambda_{x0}{}^{(m)}(\mathrm{d}^m x_i/\mathrm{d}t^m)_0 + \lambda_{y0}{}^{(m)}(\mathrm{d}^m y_i/\mathrm{d}t^m)_0+\cdots]. \tag{8}$$

The undetermined coefficients $\lambda_{x0}{}^{(0)}$, $\lambda_{x0}{}^{(1)}$, $\cdots$, $\lambda_{y0}{}^{(0)}$, $\lambda_{y0}{}^{(1)}$, $\cdots$ have values appropriate to the choice of t_0 as the time at which the derivatives $d^m x_i/\mathrm{d}t^m$ are to have reproducible values, these

[2] Here one is comparing a set of statistical analyses each of which includes a different number n of rate experiments.

[3] See R. H. Fowler, *Statistical Mechanics* (University Press, Cambridge, 1929), Section 1.5.

[4] J. W. Gibbs, *The Collected Works* (Longmans, Green andCompany, New York, 1928), Vol. 2, Part 1, Chapter 1.

coefficients being the same for all detailed behaviors.

Equation 8 expresses the frequency of occurrence of rate experiments in which the various derivatives $(\mathrm{d}^m x_i/\mathrm{d}t^m)$ will happen to have any particular set of values $(\mathrm{d}^m x_i/\mathrm{d}t^m)_0$ at the time t_0. For example, if the quantities x_i include molecular velocities, Eq. 8 expresses the distribution of velocities, of accelerations, and of higher derivatives of velocity needed at this instant in order that measurement of the quantities x_i shall give statistically reproducible results at all times. If the time t_1 had been chosen as the reference time rather than t_0, the distribution of velocities, etc., needed at the time t_1 would be given by exactly the same equation,

$$\log W_i/a_i=\lambda+\sum_{m=0}^{\infty}[\lambda_{x1}{}^{(m)}(\mathrm{d}^m x_i/\mathrm{d}t^m)_1 +\lambda_{y1}{}^{(m)}(\mathrm{d}^m y_i/\mathrm{d}t^m)_1+\cdots], \quad (9)$$

except that the values of $\lambda_x{}^{(m)}$ and $(\mathrm{d}^m x_i/\mathrm{d}t^m)$ are now those appropriate to the time t_1.

If the motion of the experimental system can be expressed by Eq. 1, then each of the derivatives $(\mathrm{d}^m x_i/\mathrm{d}t^m)_0$ may be expressed in terms of the $(\mathrm{d}^m x_i/\mathrm{d}t^m)_1$'s by a Taylor series:

$$(\mathrm{d}^m x_i/\mathrm{d}t^m)_0=\sum_{r=0}^{r=\infty}(\mathrm{d}^{m+r}x_i/\mathrm{d}t^{m+r})_1(t_0-t_1)^r/r!. \quad (10)$$

By substituting these values into Eq. 8 and collecting terms containing the same derivative, it is transformed to

$$\log W_i/a_i=\lambda+\sum_x\sum_{m=0}^{\infty}(\mathrm{d}^m x_i/\mathrm{d}t^m)_1 \times\sum_{s=0}^{s=m}\lambda_{x0}{}^{(s)}(t_0-t_1)^{m-s}/(m-s)!, \quad (11)$$

where $\sum_x$ implies summation over the different properties x, y, $\cdots$.

Equation 11 expresses the frequency of occurrence of any particular set of values $(\mathrm{d}^m x_i/\mathrm{d}t^m)$ at the time t, in a series of rate experiments which had been distributed at the time t_0 in the way required by Eq. 8, and then allowed to change in their own natural way. If the distribution law is to be more than a mathematical figment, this result must be the same as that required by Eq. 9. This will be recognized as analogous to the requirement in the Boltzmann-Enskog-Chapman[5] treatment of rate processes, that the number of molecules in any particular state of motion be stationary. However, it does not limit the possible applications of the method as severely.

Remembering that the coefficients cannot be functions of the $(\mathrm{d}^m x_i/\mathrm{d}t^m)$'s, it is evident that the coefficients of like derivatives in Eqs. 9 and 11 must be identical. Hence

$$\lambda_{x1}{}^{(m)}=\sum_{s=0}^{s=m}\lambda_{x0}{}^{(s)}(t_0-t_1)^{m-s}/(m-s)!. \quad (12)$$

Equation 12 appears to express $\lambda_{x1}{}^{(m)}$ as a function of t_0 as well as of t_1. However, since neither $\log W_i$ nor the $(\mathrm{d}^m x_i/\mathrm{d}t^m)_1$'s depend on t_0, while the $(\mathrm{d}^m x_i/\mathrm{d}t^m)_1$'s are presumed to be linearly independent in order that Eqs. 6 may all be independent, it is evident that the $\lambda_{x1}{}^{(m)}$'s can depend only on t_1 and not on t_0. Hence it must be possible to combine the various terms in Eq. 12 in such a way that t_0 vanishes.

On differentiating Eq. 12 with respect to t_1 it is found that

$$\partial\lambda_{x1}{}^{(m)}/\partial t_1=-\sum_{s=0}^{s=m-1}\lambda_{x0}{}^{(s)}(t_0-t_1)^{m-1-s}/(m-1-s)! =-\lambda_{x1}{}^{(m-1)}, \quad (13)$$

since $\lambda_{x0}{}^{(s)}$ is a function only of t_0 and not of t_1. On differentiating with respect to t_0 and collecting terms containing the same power of (t_0-t_1), one finds

$$\partial\lambda_{x1}{}^{(m)}/\partial t_0=\sum_{s=0}^{s=m}(\lambda_{x0}{}^{(s-1)}+\partial\lambda_{x0}{}^{(s)}/\partial t_0) \times(t_0-t_1)^{m-s}(m-s)!. \quad (14)$$

In collecting the terms, two summations over s were combined in one of which the first term was missing; this fact is expressed by defining

$$\lambda_{x0}{}^{(-1)}\equiv 0. \quad (15)$$

In order that $\lambda_{x1}{}^{(m)}$ be independent of t_0, it is necessary that $\partial\lambda_{x1}{}^{(m)}/\partial t_0$ be zero for all values of t_0. From Eq. 14 it is evident that this can be true only if

$$\partial\lambda_{x0}{}^{(s)}/\partial t_0=-\lambda_{x0}{}^{(s-1)}. \quad (16)$$

[5] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (The University Press, Cambridge, 1939).

Comparison of Eqs. 13 and 16, which are identical in form shows that the requirements that

(a) the $\lambda_{x0}^{(m)}$'s be independent of t_1,

(b) the $\lambda_{x1}^{(m)}$'s be independent of t_0,

(c) log W_i be independent of both t_0 and t_1, and

(d) that Eqs. 8 and 9 lead to the same result

are self-consistent.

From the foregoing it may be concluded that the general distribution law appropriate for any rate process is expressed by Eq. 8, in a form which is independent of the choice of the reference temperature provided the undetermined coefficients satisfy Eqs. 14 and 15. This distribution law is limited only in that it may not apply to processes in which the rate variables show serious discontinuities. For example, it will not apply to any molecular model assuming infinitely hard molecules, but will generally be applicable to any model in which, the molecules are more or less "soft." Thus, while the Chapman-Enskog[5] method is most suited to the study of rarefied gases, the present method is more suited to the treatment of dense phases.

SPECIAL FORMS

Rate studies in general are concerned with the transport of material and energy, and with chemical reactions which may be expressed mathematically in terms of the change of certain reaction coordinates.[6] The most detailed conditions of statistical reproducibility which might be needed to handle rate processes would thus require that the average position $\bar{\mathbf{q}}_l$, average momentum $\bar{\mathbf{p}}_l$, and average kinetic energy $(\mathbf{p}_l^2)_{Av}/2m_l$ of each elementary particle l be reproducible at all times. Furthermore, if the experimental system is isolated, the total energy E_i, will be a constant of the motion whose average value must be reproducible.

So long as the momentum $\mathbf{p}_{li}$ is expressible as $m_l \mathrm{d}\,\mathbf{q}_{li}/\mathrm{d}t$, the requirement of a reproducible average momentum is implied by the requirement of a reproducible average position, and hence is not an independent restriction. On the other hand, the average kinetic energy is not in general determined by the average momentum (e.g., $(\mathbf{p}_l^2)_{Av}$ is not equal to $(\mathbf{p}_l)_{Av}^2$, so that the requirement of a reproducible average kinetic energy is independent. Thus the distribution law for this set of restrictions is

$$\log W_i/a_i=\lambda+\lambda_E\{E_i+\sum_l\sum_{m=0}^{\infty}[\mathbf{Q}_l^{(m)}\cdot \mathrm{d}^m\mathbf{q}_{li}/dt^m+K_l^{(m)}\mathrm{d}^m\mathbf{p}_{li}^2/\mathrm{d}t^m]\}, \quad (17)$$

the summation with respect to l extending over all elementary particles.

The coefficients $\lambda_E\mathbf{Q}_l^{(m)}$ and $\lambda_E K_l^{(m)}$ are the λ's corresponding to the m time derivatives of position and momentum. They must thus satisfy Eqs. 15 and 16 in order that the form of distribution shall be invariant to a change in the reference time. Equation 17 would contain terms in the derivatives $\mathrm{d}^mE_i/\mathrm{d}t^m$ but for the fact that these derivatives are all zero so long as E_i is a constant of the motion.

Less detailed conditions of reproducibility would require only that the average position, $(\sum_l \mathbf{q}_{li}/N)_{Av}$, average total momentum, $(\sum_l \mathbf{p}_{li})_{Av}$, and average total kinetic energy $(\sum_l \mathbf{p}_{li}^2/m_l)_{Av}$ of a group of N molecules be reproducible. A restriction of this sort leads to the same result, except that $\mathbf{Q}_l^{(m)}$, and likewise $m_lK_l^{(m)}$, has the same value for all molecules of the group.

Closed Form

Equation 17 is not in a form convenient for practical calculations, since it contains an infinite number of terms. It may be reduced to a closed form by assuming that for each of the detailed behaviors, the total momentum and total kinetic energy of each group are changing at a uniform rate. This assumption will not be too severe as a first approximation, provided the number of molecules in each group is very large.

As a result of this assumption, the terms in $\mathrm{d}^m\mathbf{p}_{li}^2/\mathrm{d}t^m$ can be arranged in groups so that each group vanishes if m is 2 or greater. Similarly the terms in $\mathrm{d}^m\mathbf{q}_{li}/\mathrm{d}t^m$ may be arranged in groups all of which vanish for any values of m equal to 3 or greater. Equation 17 then reduces to

$$\log W_i/a_i=\lambda+\lambda_E\{E_i+\sum_l[\dot{\mathbf{P}}_l\cdot\mathbf{q}_{li}-\mathbf{P}_l\cdot\mathbf{p}_{li}/m_l+\mathbf{Q}_l\cdot\dot{\mathbf{p}}_{li}-\dot{K}_l\mathbf{p}_{li}^2/2m_l+K_l\mathbf{p}_{li}\cdot\dot{\mathbf{p}}_{li}/m_l]\}. \quad (18)$$

[6] Cf. Glasstone, Laidler, and Eyring, *The Theory of Rate Processes* (McGraw-Hill Book Company, New York, 1941).

In this equation it was assumed that the momentum $\mathbf{p}_{li}$ can be expressed as $m_l d\mathbf{q}_l/dt$ where m_l is a generalized mass, while $\dot{\mathbf{p}}_{li}$ is written for $d\mathbf{p}_{li}/dt$. The coefficients $\mathbf{Q}_l^{(0)}$, $\mathbf{Q}_l^{(1)}$, $\mathbf{Q}_l^{(2)}$, $K_l^{(0)}$, and $K_l^{(1)}$ of Eq. 16 have been expressed as $\dot{\mathbf{P}}_l$, $-\mathbf{P}_l$, $m_l\mathbf{Q}_l$, $\dot{K}_l/2m_l$, and $K_l/2m_l$, respectively.

Equations 15 and 16 imply that λ_E is independent of the reference time t_0. Since they also require that $\lambda_E\mathbf{Q}_l^{(0)}$ be $-\partial\lambda_E\mathbf{Q}_l^{(1)}/\partial t_0$, and that $\lambda_E\mathbf{Q}_l^{(2)}$ be $-\partial\lambda_E\mathbf{Q}_l^{(2)}/\partial t_0$, it is further implied that

$$\mathbf{P}_l = m_l d\mathbf{Q}_l/dt_0, \tag{19}$$

and

$$\dot{\mathbf{P}}_l = d\mathbf{P}_l/dt_0, \tag{20}$$

while

$$d\dot{\mathbf{P}}_l/dt_0 = 0. \tag{21}$$

In the above equations total derivatives are used since in Eqs. 15 and 16 the partial derivatives indicated simply that t_1 was being considered constant; it is not necessary to express this restriction when it is understood that the coefficients do not depend on t_1 in any way. Finally it is implied that

$$\dot{K}_l = dK_l/dt_0, \tag{22}$$

and

$$d\dot{K}_l/dt_0 = 0. \tag{23}$$

The particular form given for Eq. 18 was chosen to emphasize that $\mathbf{P}_l$ must have the dimensions of momentum. Comparison with the results given in the previous paper show that at least in some cases $\mathbf{P}_l$ is just the average particle momentum for the group to which the particle l belongs. In general, due to the appearance of the terms in $\mathbf{p}_{li}\cdot{}_{li}\dot{\mathbf{p}}$, it may be just a little different from the average momentum; for convenience call it the *characteristic* momentum of the group.

Similarly $\mathbf{Q}_l$ must have the dimensions of a coordinate (vector), while $\dot{\mathbf{P}}_l/m_l$ has the dimensions of an acceleration. Equations 19 and 20 show that $\mathbf{Q}_l$ and $\dot{\mathbf{P}}_l/m_l$ are the coordinate and acceleration which correspond to the momentum $\mathbf{P}_l$. They may thus be referred to as the characteristic coordinate and acceleration of the group, respectively, and will sometimes but not always be equal to the average coordinate and acceleration.

In the classical distribution law λ_E is evaluated[7] as $-1/kT$ where k is the Boltzmann constant and T the absolute temperature of the system. In the case of the non-homogeneous systems encountered in the study of rate processes, there may be some doubt as to the exact meaning of temperature. It is convenient, however, to write $-1/kT^0$ in place of λ_E, where T^0 is some temperature characterizing the conditions of the experiment, whose exact meaning remains to be discovered in each particular case. The coefficients K_l and $\dot{K}_l$ have no previous analogue; the first has the dimensions of time while the second is dimensionless.

Significance of the Terms

In the distribution of rate experiments among detailed behaviors which is described by Eq. 18, those detailed behaviors will occur most frequently in which each term is large positive or small negative. Since λ_E is $-1/kT^0$, and hence generally negative, each of the bracketed terms will tend toward a minimum value in the most frequent behavior.

At equilibrium only the terms λ and $\lambda_E E_i$ would appear in Eq. 18. Thus the added terms represent a displacement from the equilibrium conditions. Each term $\dot{\mathbf{P}}_l\cdot\mathbf{q}_{li}$ describes a displacement of the molecules from the equilibrium position, which will tend most frequently to make $\dot{\mathbf{P}}_l\cdot\mathbf{q}_{li}$ negative. Thus the most frequent displacement will be in a direction more or less opposite to the characteristic acceleration; the characteristic acceleration is in such a direction as to restore the molecules to their equilibrium position. The terms $\mathbf{Q}_l\cdot\dot{\mathbf{p}}_{li}$ describe the same fact in a different form, namely that the most frequent acceleration is in a direction opposite to the characteristic displacement.

The terms $-\mathbf{P}_l\cdot\mathbf{p}_{li}/m_l$ were introduced and discussed in the previous paper. There it was shown that they represent a non-Maxwellian distribution of velocity in which each group of molecules has a Maxwellian distribution about their characteristic velocity. The terms in $\mathbf{p}_{li}^2/2m_l$ may be combined with the kinetic energy part of E_i, with the result that the kinetic energy of each group is multiplied by $(1-\dot{K}_l)/kT^0$ instead

[7] Cf. J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley and Sons, New York, 1940).

of $1/kT$. In effect, each group of molecules now has a velocity distribution corresponding to a temperature $T^0/(1-\dot{K}_l)$ instead of the temperature T^0. Thus the terms $-\dot{K}_l \mathfrak{p}_{li}^2/2m_l$ allow for thermal inhomogeneities.

The terms $K_l \mathfrak{p}_{li} \cdot \dot{\mathfrak{p}}_{li}/m_l$ describe the way in which the rate of conversion of potential into kinetic energy (i.e., the rate of production $\mathrm{d}p_{li}^2/m_l \mathrm{d}t$ of kinetic energy) departs from the normal. Thus they describe frictional and collision phenomena where the previous terms described heterogeneity and driving forces. From this standpoint it is perhaps the term most important in determining such properties as viscosity, and diffusional and electric resistances.

Since the term tends as a whole to be negative, the most frequent rate of increase of kinetic energy must be opposite in sign to K_l. When the most frequent value of the kinetic energy is less than the steady state value, it must tend to increase, while if greater it must tend to decrease. Thus the sign of K_l depends on whether the most frequent value of the kinetic energy of the group is greater or less than the steady state value.

Qualitatively, the distribution law described by Eq. 18 appears to agree in major detail with the known most frequent behavior of rate experiments. Quantitatively, it can be expected to be correct only insofar as the assumption of constant rate of change of total momentum and kinetic energy is justified. This assumption is certainly justified at least in the mean for systems in which the rate processes have reached a steady state. A more complete quantitative treatment is best handled in connection with specific problems.

Other Forms

While Eqs. 17 and 18 are adequate to handle the majority of rate processes, they need not always be the most convenient. For example, the simple reaction

$$\mathrm{H}+\mathrm{H} \rightarrow \mathrm{H}_2$$

may be treated by requiring that the H−H distance have an average value which was initially large and changes reproducibly. Stated in this way, the statistical problem may be handled by means of Eq. 17 or 18. On the other hand it may be more convenient to require that the average number N_2 of molecules of H_2 be initially negligible, and increase with time in a reproducible way. In this case one must return to Eq. 8, including N_2 as one of the variables x_i, y_i, $\cdots$ to obtain a distribution law which is an extension of Gibbs' grand canonical ensemble.[8] The second of above two methods is essentially that used by Eyring[6] and co-workers in their theory of absolute reaction rates. However, their distribution law does not include terms beyond $m=0$.

At present it would be preferable in any particular problem to start with Eq. 8. A consideration of various possible ways of expressing the experimental restrictions will lead to different forms of the distribution law among which certainly one or two should be conveniently usable.

---

[8] See R. H. Fowler and E. A. Guggenheim, *Statistical Thermodynamics* (University Press, Cambridge, 1939), Chapter 6.

Letter to the Editor

THIS section will accept reports of new work, provided these are terse and contain few figures, and especially few halftone cuts. The Editorial Board will not hold itself responsible for opinions expressed by the correspondents. Contributions to this section should not exceed 600 words in length and must reach the office of the Managing Editor not later than the 15th of the month preceding that of the issue in which the letter is to appear. No proof will be sent to the authors. The usual publication charge ($3.00 per page) will not be made and no reprints will be furnished free.

The Structure of Substituted Ethylenes and Their Isomerization Polymerization and "Peroxide Addition" Reactions

ROBERT A. HARMAN AND HENRY EYRING
Princeton University, Princeton, New Jersey
July 14, 1942

THE double bond in substituted ethylenes can dissociate into a single bond in two ways.[1]

(1) By a transition from its singlet ground state (in which the two electrons have opposite spins) to a triplet state (in which the spins are the same) about 25 kcal. above the ground state—the triplet mechanism.

(2) By a rotational adiabatic transition to an upper singlet state about 40 kcal. above the ground state—the singlet mechanism.

In both cases the excited products have the great reactivity of free radicals. In case (1) the activated product will be of an atomic type because the electron spins are the same: consequently the two electrons can never occupy the same orbital on one of the carbons to give an ionic type radical but will occupy two low orbitals having little ionic character. In case (2) the electron spins are opposite; the two electrons therefore occupy the same orbital which, depending upon the other substituents, may have appreciable ionic character. In spite of the much lower activation energy required for mechanism (1), actual systems may react sometimes by one mechanism and sometimes by the other because the probability of the transition required for mechanism (1) is low.[1]

On the basis of these considerations, it is possible to correlate a considerable number of observations involving *cis-trans*-isomerization, polymerization, and "peroxide addition" reactions of ethylenic compounds. All of these reactions reveal the two types of mechanism in their behavior with catalysts and in the rate expressions for the reactions.

It would be expected[2] that paramagnetic substances, free radicals, and atoms would catalyze mechanism (1) by providing a non-homogeneous magnetic field which will act differently on each of the two magnetic dipoles arising from the spin of the two electrons of the double bond. An activated complex is thus formed with the paramagnetic substances involving the state with the spins the same, as well as the state with the spins opposed. The probability of the transition from singlet to triplet type of bond is thus increased without appreciable effect on the activation energy.

That this expected effect of paramagnetic catalysts is actually found is demonstrated by the rate expression for initiating the polymerization of styrene (a) in an atmosphere of nitrogen, (b) in the presence of benzoyl peroxide (which can react or decompose to give paramagnetic products), and (c) in the presence of free radicals.[3]

(a) $k' \quad 1.5\times10^{4}\, e^{-23300/RT}$,

(b) $k' \quad 10^{10.85}\, e^{-23500/RT}$,

(c) $k' \quad 10^{12.2}\, e^{-23500/RT}$.

The effect of paramagnetic substances including atoms in catalyzing *cis-trans*-isomerization[4,5] and "abnormal" addition[4] is likewise to be explained as a catalysis of the transition to the triplet state. The reversal of addition from the Markonikow rule when paramagnetic catalysts are used can be understood in view of the fact that the triplet activated complex is non-ionic as discussed above; consequently small differences in charge on the carbon atoms have no effect on the sense of addition of the incoming uncharged atoms or radicals, but that activated complex is favored which has the least steric hindrance and which has the greatest entropy. These effects will in general be secured when the heaviest entering group attacks the carbon with the fewest heavy groups on it (reverse of the Markonikow rule).

The second mechanism, by contrast to the first, proceeds by addition of a positively charged atom or group, and this fact accounts for catalysis of the second mechanism by the Friedel-Crafts' type catalysts in isomerization,[6] polymerizations,[7] and addition[4] reactions. The Friedel-Crafts' catalyst lowers the activation energy for forming ions by combining with the negative partner of the positive addend. The effects of resonance and induction which are found in substituted benzenes[8] operate in ethylene addition reactions by mechanism (2) to give the Markonikow addition products for *ortho-para*-directing groups and the reverse addition for *meta*-directing groups.

It is thus possible to outline a general theoretical relationship between structures of substituted ethylenes and their reactivities.

[1] J. L. Magee, W. Shand, and H. Eyring, J. Am. Chem. Soc. **63**, 677 (1941).
[2] E. Wigner, Zeits. f. physik. Chemie **B23**, 28 (1933).
[3] (a) G. V. Schulz and E. Husemann, Zeits. f. physik. Chemie **B43**, 385 (1939); (b) G. V. Schulz, H. Dinglinger, and E. Husemann, *ibid.* **B39**, 246 (1938); (c) G. V. Schulz, Naturwiss. **27**, 659 (1939).
[4] F. R. Mayo and C. Wolling, Chem. Rev. **27**, 351 (1940).
[5] B. Tamamushi and H. Akiyama, Zeits. f. Elektrochemie **43**, 156 (1937).
[6] C. C. Price and M. Meister, J. Am. Chem. Soc. **61**, 1595 (1939).
[7] H. Mark and R. Raff, *High Polymeric Reactions* (1941).
[8] T. Ri and H. Eyring, J. Chem. Phys. **8**, 433 (1940).